SCHLOSS DAGSTUHL
Leibniz-Zentrum für Informatik

Jahresbericht
Annual Report
2022

SCHLOSS DAGSTUHL
Leibniz-Zentrum für Informatik

Jahresbericht
Annual Report
2022

Herausgeber	Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH Oktavie-Allee, 66687 Wadern, Germany Telefon: +49 681 302 – 4396 Fax: +49 681 302 – 4397 E-Mail: service@dagstuhl.de
Registernummer	Amtsgericht Saarbrücken HRB 63800
Vorsitzender des Aufsichtsrates	Prof. Dr.-Ing. Dr. h. c. Stefan Jähnichen
Geschäftsführung	Prof. Raimund Seidel, Ph. D. Heike Meißner
Gesellschafter	Gesellschaft für Informatik e. V., Deutschland Universität des Saarlandes, Deutschland Technische Universität Kaiserslautern, Deutschland Karlsruher Institut für Technologie (KIT), Deutschland Technische Universität Darmstadt, Deutschland Universität Stuttgart, Deutschland Universität Trier, Deutschland Johann Wolfgang Goethe-Universität Frankfurt am Main, Deutschland Institut National de Recherche en Informatique et en Automatique (INRIA), Frankreich Centrum voor Wiskunde en Informatica (CWI), Niederlande Max-Planck-Gesellschaft zur Förderung der Wissenschaften e. V., Deutschland
Einbindung	Mitglied der Leibniz-Gemeinschaft Mitglied im Leibniz-Forschungsverbund Science 2.0
Verantwortlicher Redakteur	Prof. Raimund Seidel, Ph. D.
Redaktion	Dr. Marcel Ackermann, Dr. Andreas Dolzmann, Dr. Michael Wagner
Mitarbeit	Dr. Marcel R. Ackermann, Heike Clemens, Dr. Michael Didas, Dr. Andreas Dolzmann, Jutka Gasiorowski, Dr.-Ing. Michael Gerke, Dagmar Glaser, Oliver Hoffmann, Dr. Marsha Kleinbauer, Heike Meißner, Petra Meyer, Dr. Florian Reitz, Simone Schilke, Thomas Schillo, Christina Schwarz, Prof. Raimund Seidel, Ph. D., Dr. Michael Wagner und die Autoren der Zusammenfassungen in Kapitel 6
Gestaltung	in Zusammenarbeit mit le-tex publishing services GmbH, Leipzig, Deutschland
ISSN	2199-1995
Online-Version	https://www.dagstuhl.de/dagpub/2199-1995
Veröffentlichungsdatum	Juli 2023
Copyright	© Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH
Lizenz	Creative Commons BY-ND 4.0

Leibniz Gemeinschaft

Zertifikat seit 2012 audit berufundfamilie

Vorwort
Foreword

Nach zwei Jahren stark eingeschränktem Seminarbetrieb begann sich 2022 das LZI vom COVID-19 bedingten Stillstand zu erholen. Fast alle unserer 82 ursprünglich für 2022 geplanten Seminare fanden tatsächlich statt und nur 2 davon waren rein online und nur 16 hybrid. Die Forscherinnen und Forscher wollen unbedingt wieder nach Dagstuhl kommen, um Gleichgesinnte persönlich zu treffen und mit ihnen zusammenzuarbeiten. Die Seminarteilnahmequote ist übers Jahr stetig gewachsen und lag Ende 2022 nur mehr wenige Prozentpunkte unter den 90 Prozent, die wir typischerweise vor der Pandemie hatten.

Eine wichtige Entscheidung wurde 2022 bezüglich der Bibliothek getroffen. Der Großteil der physischen Bestände wird bestehen bleiben, aber kompakter zugänglich gemacht werden, wodurch viel Platz für andere Zwecke frei wird: einerseits für Seminarpausen, und andererseits auch für neue Büros. Der Bedarf an solchen Büroräumen ergibt sich aus der erfreulichen Tatsache, dass unsere florierenden dblp-und Verlagsabteilungen inzwischen Partner in einer Reihe erfolgreicher Großprojekte zum Aufbau von wissenschaftlichen Daten- und Informationsinfrastrukturen sind. Die Beteiligung an gleich mehreren Projektförderungen des BMBF, der DFG und auch der EU ist etwas Neues für das LZI. Wir freuen uns auf die Kooperationen.

After two years of severely restricted seminar operations, in 2022 LZI began to rebound from the near stand-still caused by the COVID-19 pandemic. Almost all of our originally 82 seminars scheduled for 2022 actually took place, with only 2 in purely on-line fashion, and only 16 in hybrid fashion. Researchers were eager to again come to Dagstuhl for in-person interaction and collaboration. The attendance rate of the seminars grew steadily and by the end of 2022 was only a few percentage points below the 90 percent that we used to see before the pandemic.

An important decision was made in 2022 regarding the library. Most of the physical holdings will remain but they will be made accessible in a more compact fashion, which will free up substantial space for other use: most for seminar breaks, but some also for new offices. The need for such office space derives from that delightful fact that our thriving bibliographic service division (dblp) and publishing division are now partners in a number of successful large grants for building up scientific data and information infrastructure. Involvement in several grants from BMBF, DFG, and also EU is something new for LZI. We look forward to the collaborations.

Im Namen der Geschäftsleitung

Prof. Raimund Seidel, Ph. D.
Wissenschaftlicher Direktor

Heike Meißner
Technisch-administrative Geschäftsführerin

On behalf of the Managing Directors

Inhaltsverzeichnis
Contents

1 Das Zentrum Schloss Dagstuhl
Schloss Dagstuhl Center

Dagstuhls Leitbild

1.1

Dagstuhl's Mission

Schloss Dagstuhl – Leibniz-Zentrum für Informatik fördert die Informatikforschung auf internationalem Spitzenniveau durch die Bereitstellung von Infrastrukturen zur wissenschaftlichen Kommunikation und für den Austausch zwischen Forschenden. Ziel von Schloss Dagstuhl ist

- die Förderung der Grundlagenforschung und der anwendungsorientierten Forschung auf dem Gebiet der Informatik,
- die wissenschaftliche Fort- und Weiterbildung im Informatikbereich,
- der Wissenstransfer zwischen Forschung und Anwendung der Informatik,
- der Betrieb einer internationalen Begegnungs- und Forschungsstätte für die Informatik.

Die Förderung und Einbindung von Nachwuchswissenschaftlern ist dabei ein wichtiger Teil dieser Aufgabe; ebenso wie der Technologietransfer zwischen Forschung und Industrie.

Schloss Dagstuhl – Leibniz-Zentrum für Informatik (Leibniz Center for Informatics) pursues its mission of furthering world class research in computer science by facilitating communication and interaction between researchers. The objective of Schloss Dagstuhl is

- to promote basic and application-oriented research in the field of informatics,
- to support advanced, scientific vocational training and to further education in the field of informatics,
- to promote the transfer of knowledge between research into informatics and application of informatics,
- and to operate an international forum and research institute for informatics.

Including and thus promoting young talents is seen as an important part of our efforts, so is promoting the exchange of knowledge and findings between academia and industry.

Entwicklung des Zentrums

Die Idee zur Gründung eines Tagungszentrum für Informatik wurde Ende der 1980er Jahre geboren, zu einem Zeitpunkt, an dem die Informatikforschung – ursprünglich der Mathematik und den Ingenieurswissenschaften entsprungen – enormen Aufwind erfuhr. Die *Gesellschaft für Informatik* beobachtete damals die zunehmende Nachfrage von Informatikwissenschaftlern am weltbekannten *Mathematischen Forschungsinstitut Oberwolfach* und sah die Notwendigkeit, ein eigens auf die Informatik ausgerichtetes Zentrum einzurichten. Schloss Dagstuhl wurde schließlich 1990 gegründet und entwickelte sich rasch zu einem weltweit renommierten Treffpunkt in der Informatikforschung. Heute beherbergt die Begegnungsstätte (siehe Fig. 1.1) normalerweise jährlich mehr als 3 000 internationale Gäste.

Seit 2005 ist Schloss Dagstuhl Mitglied in der Leibniz-Gemeinschaft, einem Verbund von 97 Forschungsinstituten, Bibliotheken und Museen.[1] Schloss Dagstuhl wird seit 2006 durch eine Bund-Länder-Förderung finanziert.

Zu dem anfänglich alleinigen Schwerpunkt des Seminarprogramms haben sich in den vergangenen Jahren zwei weitere Geschäftsfelder hinzugesellt: Zum einen der Betrieb der offenen Bibliographiedatenbank dblp, zum anderen die Angebote als Open-Access-Verleger für die Informatikforschenden.

History of the Center

The idea behind a seminar center for informatics came about during the late 1980s, when research in computer science was growing rapidly worldwide as an offshoot of mathematics and engineering. At that time the German *Gesellschaft für Informatik* (German Informatics Society) became aware of the growing number of computer scientists at the world-famous *Mathematics Research Institute* in Oberwolfach, Germany, and recognized the need for a meeting venue specific to the informatics community. Schloss Dagstuhl was founded in 1990 and quickly became established as one of the world's premier centers for informatics research. Today, Schloss Dagstuhl (see Fig. 1.1) normally hosts over 3,000 research guests each year from countries across the globe.

Since 2005, Schloss Dagstuhl has been a member of the Leibniz Association, a non-profit research consortium composed of 97 research institutes, libraries and museums throughout Germany.[1] Since 2006 the center has been jointly funded by the German federal and state governments.

Since the very first days of Schloss Dagstuhl, the seminar and workshop meeting program has always been the focus of its programmatic work. In recent years, Schloss Dagstuhl has expanded its operation and also has significant efforts underway in operating the dblp computer science bibliography and in open access publishing for the computer science community.

Seminar- und Workshop-Programm

Schwerpunkt des wissenschaftlichen Programms von Schloss Dagstuhl sind die Dagstuhl-Seminare und die Dagstuhl-Perspektiven-Workshops: Etwa 30 bzw. 45 internatio-

Seminar and Workshop Program

The Dagstuhl Seminars and Dagstuhl Perspectives Workshops form the focus of the center's work. Whereas about 30 or 45 established and young researchers gather at

[1] Stand Dezember 2022.
As of December 2022.

Fig. 1.1
Aerial photograph of Schloss Dagstuhl.

nale Forschende treffen sich eine halbe bis ganze Woche auf Schloss Dagstuhl, um im Rahmen eines Dagstuhl-Seminars intensiv über ihre aktuelle Forschung zu diskutieren. Darüber hinaus trifft sich in Dagstuhl-Perspektiven Workshops ein kleinere Gruppe von ca. 30 Forschenden der Weltspitze, um über den aktuellen Stand und die zukünftigen Schwerpunkte eines ganzen Forschungsfeldes zu beraten.

Die Seminare und Perspektiven-Workshops werden jeweils von etwa vier ausgewiesenen Forschenden im entsprechenden Gebiet beantragt. Anträge werden durch das wissenschaftliche Direktorium (siehe Kapitel 11.3) begutachtet. Stellenwert bei der Begutachtung haben neben dem eigentlichen Inhalt des Antrags auch die vorgeschlagene Gästeliste sowie die Antragsteller. Nach Annahme finden die entsprechenden Veranstaltungen dann durchschnittlich zwischen 6 und 18 Monaten später statt. Eine Teilnahme ist nur mit einer persönlichen Einladung durch das Zentrum möglich.

Das Seminarzentrum ist im und rund um das 1760 erbaute Schloss Dagstuhl beheimatet und befindet sich in einer ländlichen Gegend im nördlichen Saarland, im Herzen des Dreiländerecks Deutschland, Frankreich und Luxemburg. Es bietet den Gästen eine einzigartige Arbeitsumgebung, die den Austausch mit anderen Gästen in einer wohnlichen Atmosphäre fördert. Gemütliche Sitzecken, ansprechende Essräume, eine herausragenden Informatik-Fachbibliothek, sowie eine Vielzahl von zusätzlichen Arbeits- und Freizeiträumen bieten vielfältige Möglichkeiten, damit sich Gäste auch außerhalb des fachlichen Seminarprogramms kennenlernen und austauschen können.

Nähere Informationen über Dagstuhl-Seminare und Dagstuhl-Perspektiven-Workshops finden sich in Kapitel 2.

the Dagstuhl Seminars to report on and discuss their current work, smaller groups of about 30 of the international elite of a field gather at the Dagstuhl Perspectives Workshops for the purpose of reflecting on the current status of research and potential development perspectives.

The Dagstuhl Seminars and Perspectives Workshops are characterized by the fact that they are subject to an exacting quality assurance process. A small group of up to four scientists of international standing submit a proposal for a seminar on a specific research topic. The proposal is reviewed by the center's Scientific Directorate (see Section 11.3) with regard to its content, the proposed guest list and those submitting the proposal. The seminars and workshops are held 6 to 18 months later in the seclusion of the center's facilities at Dagstuhl Castle. Participation in a seminar is possible only by way of personal invitation by the center.

Located in a 1760 built manor house in the idyllic countryside of northern Saarland at the heart of the tri-country region formed by Germany, France and Luxembourg, Schloss Dagstuhl offers visitors a unique working environment that encourages guests to interact with each other in tandem with daily life. Lounges, formal and informal dining areas, a world-class research library, and an impressive range of work and leisure rooms offer multiple possibilities for connecting one-on-one outside of the official conference rooms and meeting times.

More information on the Dagstuhl Seminars and Dagstuhl Perspectives Workshops can be found in Chapter 2.

Bibliographiedatenbank dblp

Bereits seit 2011 betreibt Schloss Dagstuhl in enger Zusammenarbeit mit der Universität Trier die Bibliographiedatenbank dblp. Seit November 2018 ist Schloss Dagstuhl in vollem Umfang alleine für den Betrieb der Datenbank verantwortlich.

Mit mittlerweile mehr als 6,4 Millionen Publikationseinträgen ist dblp die weltweit größte offene Sammlung bibliographischer Daten in der Informatik. Der dblp-Dienst ist darauf ausgerichtet, Forschenden bei ihrer täglichen Arbeit zu unterstützen, etwa bei der Literaturrecherche oder beim Bezug von elektronisch verfügbaren Volltexten. Dabei gilt dblp in der Informatik insbesondere als die Referenzdatenbank für qualitätsgesicherte, normierte Bibliographiedaten. Aber auch Forschungsförderer und Entscheidungsträger unterstützt dblp, etwa durch das Pflegen und öffentlich Verfügbarmachen von personalisierten Publikationsnachweisen. Durch den Betrieb von dblp leistet Schloss Dagstuhl einen weiteren Beitrag im Rahmen seiner Mission zur Förderung der Erkennung, Verbreitung und Umsetzung neuer Informatikerkenntnisse auf internationalem anerkanntem Niveau.

Details über dblp finden sich in Kapitel 3.

Dagstuhl Publishing

Die Förderung der Kommunikation zwischen den Wissenschaftlern in der Informatik gehört zu der zentralen Aufgabe von Schloss Dagstuhl. Wissenschaftliche Veröffentlichungen sind Teil der Forschungskultur, um qualitätsgesicherte Forschungsergebnisse zu diskutieren und zu kommunizieren. Mit seinen Open-Access-Verlagsangeboten unterstützt Schloss Dagstuhl die Forschungsgemeinde dabei, freien Zugang zu den wichtigsten und neuesten Forschungsergebnissen zu erlangen.

Neben Veröffentlichungen, die in engem Bezug zum wissenschaftlichen Programm stehen, verlegt Schloss Dagstuhl auch Konferenzbände und Zeitschriften. Herausragende Reihe ist dabei LIPIcs, in der die Publikationen erstklassiger Konferenzen erscheinen. Alle Angebote der Verlagsabteilung werden durch international besetzte Editorial Boards qualitätsgesichert.

Kapitel 4 stellt Dagstuhls Verlagswesen ausführlicher dar.

dblp computer science bibliography

Since 2011, Schloss Dagstuhl has been operating the dblp computer science bibliography in close cooperation with the University of Trier. In November 2018, Schloss Dagstuhl alone assumed full responsibility for the operation of the database.

Listing more than 6.4 million articles, dblp is the world's most comprehensive open data collection of computer science research articles. The goal of dblp is to support computer scientists in their daily work, for example when reviewing the literature of a given author or subject area, or when searching for online full-text versions of research articles. The dblp database is often considered to be the reference database for quality-assured and normalized bibliographic metadata in computer science. Additionally, dblp supports funding agencies and decision makers by providing and curating personalized author bibliographies. By operating dblp, Schloss Dagstuhl furthers its mission of promoting the identification, dissemination and implementation of new computer science developments at an internationally recognized level.

More information about the dblp computer science bibliography can be found in Chapter 3.

Dagstuhl Publishing

Enabling communication between researchers in computer science is part of Dagstuhl's central mission. Scholarly publications belong to the culture of discussing and communicating quality-controlled research results on a global level. Dagstuhl's open-access publishing services hence support the need of the research community to have access to the most important and most recent research results.

In addition to the open documentation of proceedings of its seminar and workshop program, Schloss Dagstuhl also publishes proceedings for computer science conferences and journals. The flagship product of Dagstuhl Publishing is the LIPIcs series, which publishes proceedings of outstanding computer science conferences. The scientific quality of all products is supervised by international editorial boards.

More information on Dagstuhl Publishing can be found in Chapter 4.

Neuigkeiten in 2022 — 1.2 — *News from 2022*

Pandemie

Im Laufe des Jahres 2022 normalisierten sich die Abläufe im Bereich des Seminarwesens. Pandemiebedingte Einschränkungen konnten unter dem Schutz regelmäßiger Schnelltests schrittweise fallen gelassen werden und bis auf kleinere Einschränkungen und einer etwas geringeren Anzahl von Teilnehmenden, war das Veranstaltungsprogramm Ende des Jahres fast wieder auf dem Niveau vor der Pandemie. Näheres zum Seminarwesen und den Einschränkungen dort finden sich im Abschnitt „Seminare und

Pandemic

Over the course of 2022, seminar related processes returned to normal. Using the protection provided by regular rapid tests, pandemic-related restrictions were gradually removed and, apart from minor restrictions and a slightly lower number of participants, the event program was almost back to its pre-pandemic level by the end of the year. More details about the seminar process and the restrictions affecting it are given in section "Seminare und Workshops" below. After a two-year break, the teacher training and

Workshops“ weiter unten. Nach zwei Jahren Pause konnte endlich wieder die Lehrerfortbildung und auch die Kooperation mit dem Heidelberg Laureate Forum in gewohnter Form stattfinden. Die Reihe „Dagstuhler Gespräche“, die noch pausiert, soll bald wieder fortgesetzt werden. Der Bereich Publishing war bis auf wenige Terminverschiebungen von Konferenzen kaum betroffen, der Bereich dblp gar nicht.

32 Jahre Informatik in Schloss Dagstuhl

Der 2020 wegen der Pandemie ausgefallene Festakt anlässlich 30 Jahre Informatik in Schloss Dagstuhl wurde als „2^5 Years of Computer Science at Schloss Dagstuhl“ am 24 Juni 2022 nachgeholt. An das Festkolloquium mit Grußworten aus der Politik und wissenschaftlichen Vorträgen schloss sich ein Grillempfang an. Details und das Programm des Kolloquiums findet sich in Kapitel 7.

Das Team

Am Ende des Jahres 2022 beschäftigte Schloss Dagstuhl insgesamt 41,09 Vollzeitäquivalente bzw. 57 Angestellte.

Das LZI hat in 2022 eine Mitarbeiterin im Wissenschaftlichen Stab im Bereich Seminarwesen und 4 Mitarbeiter für die neu gestarteten Projekte neu eingestellt. Weiterhin wurde das Team für Küche und Hausdienst mit 4 Mitarbeiterinnen verstärkt. 2 neue Kolleginnen übernehmen den Dienst an der Rezeption am Sonntag.

Schloss Dagstuhl freut sich, dass eine Auszubildende die Landesbeste beim Landeswettbewerb der Hauswirtschafter wurde. Sie darf nun am Bundeswettbewerb der Hauswirtschafter im März 2023 in Miesbach teilnehmen.

Nationale Forschungsdateninfrastruktur (NFDI)

Beginnend in den Jahren 2022 und 2023 hat das LZI Schloss Dagstuhl Fördermittel für die Teilnahme an zwei NFDI-Konsortien erhalten. Die NFDI ist ein kollaboratives, bundesweites Netzwerk, um den wertvollen Datenbestand aus Wissenschaft und Forschung systematisch zu erschließen, zu vernetzen und offen zugänglich zu machen. Gemeinsam mit Dutzenden anderer führender Partnerinstitutionen will Schloss Dagstuhl Forschungsdaten und Forschungsartefakte aus den Bereichen Informatik und Datenwissenschaften systematisch verwalten:

- NFDI for and with Computer Science (NFDIxCS).[2] Das Hauptziel von NFDIxCS ist es, zentrale Infrastrukturdienste für die Informatik-Community zu definieren und bereitzustellen und die FAIR-Prinzipien (Auffindbarkeit, Zugänglichkeit, Interoperabilität und Wiederverwendbarkeit) flächendeckend in unserer Disziplin umzusetzen. Dazu gehört die Erstellung wiederverwendbarer digitaler FAIR-Objekte in Form von *Research Data Management Containers (RDMCs)*, die

also the cooperation with the Heidelberg Laureate Forum could finally take place again in the usual format. The series “Dagstuhler Gespräche” (Dagstuhl Talks), is still currently on hold, but should fully resume again in the near future. Apart from a few postponements of conferences, the publishing department was hardly affected, and the dblp department was not at all affected.

32 Years Informatics in Schloss Dagstuhl

The ceremony for the anniversary of 30 years of computer science at Schloss Dagstuhl, which was cancelled in 2020 due to the pandemic, was made up for with the Scientific Colloquium and Anniversary Ceremony "2^5 Years of Computer Science at Schloss Dagstuhl" on June 24, 2022. The festive colloquium with welcoming addresses from politicians and with scientific lectures was followed by a barbecue reception. Details and the programme of the colloquium can be found in chapter 7.

The Team

By the end of 2022, Schloss Dagstuhl had a total of 57 staff members corresponding to 41.09 full-time positions.

In 2022, Schloss Dagstuhl LZI newly hired one employee for the scientific staff in the area of seminar preparation and 4 employees for newly started projects. Furthermore, the kitchen and housekeeping team was strengthened with 4 new employees. In addition, 2 new colleagues have been hired to run the reception service on Sundays.

Schloss Dagstuhl is pleased to announce that one of its trainees won a state competition for home economics. She is now allowed to participate in the national competition for home economics in March 2023 in Miesbach.

National Research Data Infrastructure NFDI

Starting in 2022 and 2023, Schloss Dagstuhl LZI has been awarded funding to join two NFDI consortia. The NFDI is a collaborative, nationwide network to systematically index, interconnect, and make openly available the valuable stock of data from science and research. Together with dozens of other leading partner institutions, Schloss Dagstuhl aims to systematically manage research data and artifacts from the computer science and data science communities:

- NFDI for and with Computer Science (NFDIxCS).[2] The main goal of NFDIxCS is to define and deploy central infrastructure services for the computer science community, and to implement the *FAIR principles* (findabily, accesibility, interoperability, and reuse) across the board in our discipline. This includes producing reusable FAIR digital objects in the form *Research Data Management Containers (RDMCs)* which contain not only various domain-specific types of data and their associated metadata, but also the corresponding

nicht nur verschiedene domänenspezifische Datentypen und die zugehörigen Metadaten enthalten, sondern auch die entsprechende Software-, Kontext- und Ausführungsinformationen in standardisierter Form.

- NFDI for Data Science and AI (NFDI4DataScience).[3] Die Vision von NFDI4DataScience ist es, alle Schritte des komplexen und interdisziplinären Lebenszyklus von Forschungsdaten zu unterstützen, einschließlich der Sammlung, Erstellung, Verarbeitung, Analyse, Veröffentlichung, Archivierung und Wiederverwendung von Ressourcen in den Bereichen Datenwissenschaft und künstliche Intelligenz. Digitale FAIR-Objekte werden mit reichhaltigen Metadateninformationen dargestellt und in einem gemeinsamen Research Knowledge Graph (RKG) zusammengeführt, der ein transparentes Verständnis der Code-/Datenherkunft, der Modellkonfigurationen und der Trainings-/Testdaten ermöglicht.

Unsere international anerkannten Forschungsinformationsinfrastrukturen – die Open-Access-Publikationsplattform DROPS und die dblp-Informatikbibliographie – werden eine entscheidende Rolle bei der Erfüllung dieser Aufgabe spielen. Das Konsortium wird von Bund und Ländern für einen Zeitraum von zunächst 5 Jahren finanziert. Die Ergebnisse der NFDI werden von unseren Infrastrukturen für die Gemeinschaft als offene Dienste, offene Forschungsdaten und offene Wissensgraphen weit über die anfängliche Förderdauer hinaus gepflegt und erweitert werden.

FAIRCORE4EOSC

Seit dem Juni 2022 beteiligt sich Schloss Dagstuhl mit seiner Abteilung Dagstuhl Publishing am Projekt FAIRCORE4EOSC[4], welches sich auf die Entwicklung und Realisierung von Kernkomponenten für die European Open Science Cloud (EOSC) konzentriert und durch die europäische Union gefördert wird. Schwerpunkt der Beteiligung des LZI liegt dabei in der Integration von Werkzeugen zur Archivierung, Referenzierung, Beschreibung und Zitierung von wissenschaftlicher Software in den Veröffentlichungsprozess von wissenschaftlichen Arbeiten. Der Förderzeitraum dauert von Juni 2022 bis Mai 2025.

Seminare und Workshops

Die fortdauernde Covid-19 Pandemie beeinflusste auch im Jahr 2022 das Seminarwesen, das dennoch von einer Rückkehr zur Normalität geprägt war.

Mit 77 Dagstuhl-Seminaren und Dagstuhl-Perspektiven-Workshops – hier und im folgenden unter „Seminare“ subsumiert – war das Zentrum überdurchschnittlich gut belegt. Nur 5 der eigentlich geplant 82 Seminaren wurden pandemie-bedingt abgesagt, nur 2 fanden rein online statt und nur noch 16 hybrid.

Verglichen mit 2019 hatten Schloss Dagstuhls Seminare, obwohl 4 Seminare mehr stattfanden, in der Summe etwa 10 % weniger Teilnehmende und etwa 25 % Teilnehmende bzw Übernachtungen weniger vor Ort.

software, context and execution information in a standardized form.

- NFDI for Data Science and AI (NFDI4DataScience).[3] The vision of NFDI4DataScience is to support all steps of the complex and interdisciplinary research data lifecycle, including collecting, creating, processing, analyzing, publishing, archiving and reusing resources in Data Science and Artificial Intelligence. FAIR digital objects will be represented with rich metadata information and assembled into a joint *Research Knowledge Graph (RKG)*, enabling a transparent understanding of code/data provenance, model configurations, and training/testing data.

Our internationally renown research information infrastructures – the Open Access publishing platform DROPS and the dblp computer science bibliography – will play a crucial role in fulfilling this mission. The consortium is funded by the federal and state governments for an initial period of 5 years. Results of NFDI will be maintained and extended by our infrastructures for the community as open services, open research data, and open knowledge graphs well beyond the initial funding duration.

FAIRCORE4EOSC

Since June 2022, Schloss Dagstuhl with its division Dagstuhl Publishing has been participating in the project FAIRCORE4EOSC[4], which focuses on the development and realisation of core components for the European Open Science Cloud (EOSC) and is funded by the European Union. The focus of LZI's participation is on the integration of tools for archiving, referencing, describing and citing scientific software in the publication process of scientific papers. The funding period runs from June 2022 to May 2025.

Seminars and Workshops

The ongoing Covid-19 pandemic continued to influence the seminar program in 2022, which was nevertheless characterized by a return to normality.

With 77 Dagstuhl Seminars and Dagstuhl Perspectives Workshops – from here on simply referred to as "seminars" – the center had above-average attendance. Only 5 of the 82 actually planned seminars were canceled due to the pandemic, just 2 took place purely online and just 16 in a hybrid format.

Compared to 2019, Schloss Dagstuhl's seminars had about 10 % fewer participants and about 25 % fewer attendees or overnight stays on site, despite 4 more seminars.

[2] https://nfdixcs.org/
[3] https://www.nfdi4datascience.de/
[4] https://faircore4eosc.eu/

Die vor-Ort Auslastung der Seminare, also das Verhältnis von Plätzen zu Teilnehmenden stieg 2022 von 48 % im ersten Quartal auf 82 % im vierten Quartal – 2019 lag sie knapp bei konstanten 90 %.

Mit 93 Anträgen auf Dagstuhl-Seminare und Dagstuhl-Perspektiven Workshops wurden etwas weniger Veranstaltungen als das langjährige Mittel von 100 Anträgen pro Jahr beantragt. Dies mag auch daran liegen, dass das Seminarprogramm von Schloss pandemiebedingt aktuell längere Vorlaufszeiten hat. Trotz allem sichern diese aber eine gute Auslastung des Zentrums für 2023 und teils für 2024.

Mehr Details und Zahlen zum Seminarprogramm finden sich in Kapitel 2.

dblp als Forschungs-Wissensgraph

Seit mehr als 20 Jahren wird ein vollständiger, täglich aktualisierter XML-Dump der dblp-Bibliografie der Informatik als offene Daten zum Herunterladen und zur Wiederverwendung zur Verfügung gestellt. Schnappschüsse des XML-Dump wurden bereits von Mitgliedern der Community in RDF konvertiert. Diese Schnappschüsse sind jedoch in der Regel nicht mit den kontinuierlich kuratierten dblp-Daten synchronisiert, in einigen Fällen sogar mehrere Jahre veraltet. Um dieses Problem zu beheben, hat das dblp-Team Anfang 2022 damit begonnen, seine Daten auch als semantische Daten in RDF über APIs und als full dump download.[5] zu veröffentlichen.

Unser Ziel ist es, die vielfältigen Informationen, die von der dblp-Website (wenn auch oft nur unstrukturiert) bereitgestellt werden, in einen semantisch reichhaltigen, maschinenverarbeitbaren Wissensgraphen für die Forschung zu verwandeln. Dieser Wissensgraph ist auf dem neuesten Stand und synchron mit dem kontinuierlich kuratierten und disambiguierten dblp Datenbestand. Wie alle anderen von dblp bereitgestellten Daten werden auch die RDF-Daten unter der CC0-Lizenz zur Verfügung gestellt.

In seiner ersten Version bildet der dblp-Wissensgraph einen einfachen Person-Publikations-Graphen, der (Stand Januar 2023) aus insgesamt 347 Millionen RDF-Tripeln besteht. Zahlreiche Metadatenaspekte, wie z. B. Zeitschriften/Konferenzreihen oder die Zugehörigkeit eines Autors, werden derzeit noch als Stringliterale bereitgestellt. In zukünftigen Iterationen des Graphen werden diese und weitere Aspekte als echte Entitäten mit eigenen Metadaten, persistenten IDs und Links zu externen Ressourcen hinzugefügt werden. Daher sehen wir den aktuellen dblp-Wissensgraphen nicht als endgültig an, sondern eher als einen ersten Schritt, um die Semantik des dblp-Datensatzes auf strukturiertere Weise bereitzustellen. Wir werden in naher Zukunft auch einen geeigneten SPARQL-Endpunkt bereitstellen.

Dagstuhl Publishing

Wie in den Vorjahren haben die Open-Access-Publikationsaktivitäten auch in 2022 starken Zuspruch bekommen. So wurden in den Konferenzbandreihen LIPIcs und OASIcs

On-site seminar occupancy, i.e., the ratio of places to participants, increased in 2022 from 48 % in the first quarter to 82 % in the fourth quarter – but in 2019 it remained constant at about 90 %.

With 93 applications for seminars, slightly fewer events were requested than the long-term average of 100 applications per year. This may also be due to the fact that the Schloss Dagstuhl seminar program currently has longer lead times as a result of the pandemic – but in any case, this will ensure good utilization of the center's capacity for 2023 and partly for 2024.

See Chapter 2 for more details and statistics regarding the seminar program.

dblp as a research knowledge graph

For more than 20 years, a full, daily updated XML dump of the dblp computer science bibliography has been made available as open data for download and reuse. Snapshots of the dblp XML dump have been converted to RDF before by members of the community. However, these snapshots are usually severely out of sync with the continuously curated dblp data, in some cases up to several years. To remedy this problem, the dblp team has begun in early 2022 to release its data also as semantic data in RDF via APIs and as a full dump download.[5]

Our goal is to transform the wide range of diverse information provided by the dblp website (although often only in an unstructured way) as a semantically rich, machine-actionable research knowledge graph. This knowledge graph is up to date and in sync with the continuously curated and disambiguated dblp data corpus. Just as with any other data provided by dblp, RDF data is made available under CC0 license.

In its initial release, the dblp knowledge graph forms a simple person-publication graph, consisting (as of January 2023) of 347 million RDF triples in total. Numerous metadata aspects, like journals/conference series or the affiliation of an author, are currently still provided only as string literals. Future iterations of the graph will see these and further aspects being added as true entities, together with their own metadata, persistent IDs, and links to external resources. Hence, we do not see the current dblp knowledge graph as final, but rather as a first step in providing the semantics of the dblp dataset in a more structured way. We will also provide a proper SPARQL endpoint in the near future.

Dagstuhl Publishing

As in the previous years, Schloss Dagstuhl's open-access publishing services experienced an on-going strong increase in demand from the community in 2022. So

[5] https://dblp.org/rdf/

zusammen 1 539 Publikationen innerhalb eines Jahres veröffentlicht.

Zudem gab es auch in 2022 wieder viele Bewerbungen von wissenschaftlichen Konferenzen zur Veröffentlichung des Konferenzbandes in den Serien LIPIcs und OASIcs. Bei LIPIcs wurde in 10 Fällen ein bereits bestehender Vertrag (auf Antrag der jeweiligen Konferenz und nach intensiver Prüfung durch das Editorial Board) um (in der Regel) 5 Jahre verlängert. Zusätzlich wurde eine Konferenz für 5 Jahre neu ins Portfolio von LIPIcs aufgenommen.

In der wissenschaftlichen Zeitschrift LITES wurden im Berichtsjahr zwei Sonderhefte mit insgesamt 10 Artikeln veröffentlicht. Zudem wurde mit Björn B. Brandenburg der Nachfolger des bisherigen Hauptherausgebers Alan Burns bestimmt. Dieser wird seine Tätigkeit Anfang 2023 aufnehmen. Wir danken Alan Burns auf diesem Wege herzlich für die hervorragende Zusammenarbeit und freuen und auf die Zusammenarbeit mit Björn B. Brandenburg.

Mehr Informationen zu den Open-Access-Aktivitäten von Schloss Dagstuhl finden sich in Kapitel 4.

Lehrerfortbildung

In Zusammenarbeit mit dem saarländischen Landesinstitut für Pädagogik und Medien (LPM) und dem Pädagogischen Landesinstitut Rheinland-Pfalz (PL) organisierte Schloss Dagstuhl 2022 zum 30. Mal eine Lehrerfortbildung, die sich an Informatik- und Mathematiklehrer der gymnasialen Oberstufe im Saarland und in Rheinland-Pfalz richtet. Dem Jubiläum der 30. Lehrerfortbildung wurde mit einem Grußwort der saarländischen Ministerin für Bildung und Kultur, Christine Streichert-Clivot, und einem Videogrußwort der rheinland-pfälzischen Ministerin für Bildung, Dr. Stefanie Hubig gedacht.

Mehr Informationen zur Öffentlichkeitsarbeit und zu den Weiterbildungsaktivitäten finden sich in Kapitel 7.

Zusammenarbeit mit dem Heidelberg Laureate Forum

Nach der erzwungenen Pause durch die Pandemie gab es 2022 wieder eine Kooperation von Schloss Dagstuhl mit dem Heidelberg Laureate Forum[6] (HLF). Diese Veranstaltung bringt herausragende Forscher, die mit dem ACM Turing Award, dem Abelpreis, der Fields-Medaille oder mit dem Nevanlinna-Preises ausgezeichnet wurden, mit außergewöhnlich begabten jungen Wissenschaftlern aus aller Welt zusammen. Zwei ausgewählte Teilnehmende des HLF 2022 erhielten in der Woche nach der neunten Ausgabe dieses Forums die Gelegenheit zur Teilnahme an dem Dagstuhl-Seminar „Transparent Quantitative Research as a User Interface Problem“ (22392), ein Weiterer konnte am Dagstuhl-Seminar „Cognitive Robotics“ (22391) teilnehmen. Aufgrund des großen Erfolgs der Initiative haben alle Partner einer Fortsetzung der Zusammenarbeit für das Jahr 2023 zugestimmt.

in the conference proceedings series LIPIcs and OASIcs together, 1,539 publications were published within one year.

In addition, there were again many applications from scientific conferences for publication of the conference proceedings in the LIPIcs and OASIcs series in 2022. As far as LIPIcs is concerned, in 10 cases an already existing contract was extended for (usually) 5 years (at the request of the respective conference and after intensive review by the Editorial Board). In addition, one conference was newly included in the LIPIcs portfolio for 5 years.

In the scientific journal LITES two special issues with a total of 10 articles were published in the reporting year. In addition, Björn B. Brandenburg was appointed as the successor to the previous Editor-in-Chief Alan Burns. He will take up his position at the beginning of 2023. We would like to take this opportunity to thank Alan Burns for his excellent cooperation and look forward to working with Björn B. Brandenburg.

More information about the Open Access activities of Schloss Dagstuhl can be found in Chapter 4.

Teacher Training Program

In 2022, Schloss Dagstuhl hosted its teacher training course for the 30th time. This workshop is specifically designed for computer science and mathematics teachers teaching grades 11 and 12 in Saarland and Rhineland-Palatinate. It is organized in collaboration with the Landesinstitut für Pädagogik und Medien Saarland LPM (Saarland State Institute for Education and Media) and the Pädagogisches Landesinstitut Rheinland-Pfalz PL (Rhineland-Palatinate State Institute for Education). The 30th anniversary of the teacher training program was commemorated with a greeting from the Saarland Minister for Education and Culture, Christine Streichert-Clivot, and a video greeting from the Rhineland-Palatinate Minister for Education, Dr. Stefanie Hubig.

Further details about public relations and professional training at Schloss Dagstuhl can be found in Chapter 7.

Joint Outreach with the Heidelberg Laureate Forum

After the involuntary break forced by the pandemic, 2022 saw another cooperation venture between Schloss Dagstuhl and the Heidelberg Laureate Forum[6] (HLF). The HLF brings winners of the ACM Turing Award, the Abel Prize, the Fields Medal, and the Nevanlinna Prize together with exceptionally talented young scientists from all over the world. Two participants were selected and participated in the Dagstuhl Seminar “Transparent Quantitative Research as a User Interface Problem” (22392), taking place during the week after the ninth edition of the forum. Another participant was selected and participated in the Dagstuhl Seminar “Cognitive Robotics” (22391). Satisfied with the outstanding success of the initiative, all partners have agreed to continue the cooperation in 2023.

[6] http://www.heidelberg-laureate-forum.org

Spender und Förderer

Schloss Dagstuhl ist den wissenschaftlichen Gästen, Institutionen und Firmen dankbar, die durch großzügige Spenden das Zentrum unterstützen.

2022 erhielt die Bibliothek von Verlagshäusern erneut zahlreiche Buchspenden. Insgesamt erhielt das Zentrum im Berichtszeitraum 569 Bände als Spende, darunter 553 Monographien des Springer-Verlags im Wert von 45 406,85 €.

Schloss Dagstuhl erhielt 2022 verschiedene Kunstspenden. Schloss Dangstuhl dankt Maria Krause[7] für die Schenkung von drei Skulpturen, die nun im Schloss ausgestellt werden. Ebenso dankt Schloss Dagstuhl Hsiang-Yun Wu für ihr Werk „The World of Scientists".

Neuer Webauftritt

Ende 2022 ging ein von Mitarbeitern komplett neu entwickelte Webseite mit einem modernen und frischem Aussehen für Schloss Dagstuhls Internetauftritt online. Nach 16 Jahren trägt der Webauftritt mit einer angepassten und verschlankten Seitenstruktur dem über die Jahre gewachsenen Leistungsspektrum von Schloss Dagstuhl Rechnung. Die Umstellung auf ein aktuelles Contentmanagementsystem und Web-Framework führt zu einer barrierearmen, responsiven und smartphonefreundlichen Webpräsenz.

Baumaßnahmen und Renovierung

Neben den üblichen Malerarbeiten und Reparaturen in der Sommerpause gab es im Jahr 2022 viele Veränderungen: Es wurde am Haupteingang des Zentrums eine Säule mit Klingel, Kartenleser und Zahleneingabeblock errichtet, damit der Zugang zum Gebäude durch die automatischen Schiebetüren für Rollstuhlfahrer leichter ermöglicht wird.

Der sogenannte Neubau wurde zusätzlich mit einem Gasheizungsboiler versehen, bisher wurde dort ausschließlich mit Öl geheizt. Zur Senkung der Energiekosten wurden SmartHome Heizkörperthermostaten in allen öffentlichen Bereichen angebracht, und ein Großteil der Beleuchtung wurde sukzessive auf LEDs umgestellt.

Für eine solide Planung und Umsetzung einer Nachhaltigkeitsstrategie hat Schloss Dagstuhl die Erstellung eines energetischen Gesamtkonzeptes beauftragt und auf dessen Basis einen Stufenplan erstellt. Umsetzungen erster Maßnahmen sind für das Jahr 2023 geplant. Zu unseren Nachhaltigkeitsmaßnahmen gehören auch zwei Regenwassertanks im Gästehaus sowie die Umnutzung eines alten gereinigten Öltanks, um Regenwasser zur Bewässerung der Liegenschaften zu sammeln.

Ausstattung

Im Konferenzsaal „Saarbrücken" werden, aufgrund der Rückmeldungen unserer Gäste zu ihren Erfahrungen mit Tischen statt reiner Bestuhlung während der Pandemie, die

Sponsors and Donors

Schloss Dagstuhl is grateful to its scientific guests and institutional colleagues for generous donations for the support of its center.

In 2022, the center's research library again received a large number of book donations from publishing houses. The number of donated volumes totaled 569, including 553 monographs at a total value of 45,406.85 € donated by Springer-Verlag.

Schloss Dagstuhl received various art donations in 2022. Schloss Dagstuhl would like to thank Maria Krause[7] for the donation of three sculptures, which are now on display in the castle. Likewise, Schloss Dagstuhl would like to thank Hsiang-Yun Wu for her art piece: "The World of Scientists".

New Website

At the end of 2022, a website completely redeveloped by staff went online with a fresh and modern look for Schloss Dagstuhl's presence on the internet. After 16 years, the website showcases Schloss Dagstuhl's range of services, which has grown substantially over the years, by the use of an adapted and streamlined page structure. This conversion to an up-to-date content management system and web framework has resulted in a reduced-barrier, responsive, and smartphone-friendly web presence.

Construction Work and Renovation

Aside from the necessary paintwork and small repairs during the vacation close-down, 2022 saw a lot of changes: a column with a doorbell, a card reader, and a key-pad was erected at the main entrance of the centre to make it easier for wheelchair users to access the building through the automatic sliding doors.

The so-called new building was equipped with an additional gas heater and boiler; previously, it was heated exclusively with oil. To reduce energy costs, SmartHome radiator thermostats were installed in all public areas, and much of the lighting was successively converted to LEDs.

For solid planning and implementation of a sustainability strategy, Schloss Dagstuhl has commissioned the creation of an overall energy concept and, based on this, has drawn up a step-by-step plan. Implementation of the first measures is planned for 2023. Our sustainability measures also include two rainwater tanks in the guest house as well as the conversion of an old cleaned oil tank to collect rainwater for irrigation of the properties.

Facilities

Based on feedback from our guests about their experiences over the course of the pandemic with tables instead of just chairs, the lecture hall "Saarbrücken" will now

[7] http://maria-krause.de/

Stühle jetzt dauerhaft um kleinere Konferenztische ergänzt.

Der Raum „Kaiserslautern“ wurde mit neuen Komponenten zur Projektion und für hybride Veranstaltungen ausgestattet, so dass er nun den beiden anderen großen Hörsälen ebenbürtig ausgestattet ist.

In Zusammenarbeit mit dem Energieversorger Energis wurde auf dem Parkplatz von Schloss Dagstuhl eine öffentliche Ladesäule mit zwei Ladepunkten für Elektroautos gebaut.

Seit neustem hält Schloss Dagstuhl einen Aufenthaltsraum für Kinder vor. Dieser ist unter anderem mit einer Sitzgruppe, einem Spielhaus, einer Spielhöhle und Kinderzimmerraumteilern, die etwa mit einer Kugelbahn oder einem Spiegel ausgestattet sind, versehen. Er erlaubt einen angenehm Aufenthalt der Kinder auch im Innenraum.

Bibliothek

Zu den meisten Dagstuhl-Seminaren stellt Schloss Dagstuhl alle in der Bibliothek vorhandenen Bücher der anwesenden Gäste zusammen und präsentiert sie in einer separaten Auslage. Durchschnittlich werden mehr mehr als 50 Bücher pro Dagstuhl-Seminar präsentiert.

Die umfangreiche Sammlung „Dagstuhls Impact“[8], die publizierte Ergebnisse aus Dagstuhl-Seminaren dokumentiert, wurde 2022 um zahlreiche weitere Publikationen ergänzt: 82 Artikel, 3 Zeitschriftensonderhefte und 3 Bücher konnten hinzugefügt werden.

Aus Platzgründen wird es künftig eine fahrbare Regalanlage im Bibliothekskeller geben. Nach einer intensiven Planungsphase konnte mit den vorbereitenden Arbeiten dazu begonnen werden. Der komplette Buchbestand des Bibliothekskellers wurde nach einer Bestandsrevision hausintern inklusive aller Regale umgezogen, so dass Platz für den Aufbau der neuen Regalanlage geschaffen wurde, die Anfang 2023 aufgebaut werden soll.

Kunst

Seit 1995 finden in Schloss Dagstuhl regelmäßig Kunstausstellungen statt. Bisher wurden die Ausstellungen meist von einem einzelnen Künstler oder manchmal auch von einer kleinen Gruppe von Künstlern gestaltet. Im Berichtsjahr wurde nun eine Zusammenarbeit zwischen dem Institut für aktuelle Kunst im Saarland und Schloss Dagstuhl vereinbart. Als erste Ausstellung dieser Zusammenarbeit entstand „THINKING STRUCTURES“ mit Arbeiten von Monika von Boch, Jo Enzweiler und Vera Molnár. Dankenswerterweise unterstützt der Verein zur Förderung von Schloss Dagstuhl – Leibniz-Zentrum für Informatik e.V. die Zusammenarbeit.

Weitere Informationen über das generelle Konzept, die Ausstellung in 2022 und die neugestartete Zusammenarbeit finden sich in Kapitel 10.

permanently add smaller conference tables to its seating arrangement.

The hall “Kaiserslautern” was equipped with new components for projection and hybrid events, so that its equipment is now comparable to the other two large lecture halls.

In cooperation with the energy provider Energis, a public charging station with two charging points for electric cars was added to the parking lot of Schloss Dagstuhl.

Schloss Dagstuhl has recently added a recreation room for children. It includes a seating area, a playhouse, a play cave, and children’s room dividers equipped with a marble run, a mirror, and so on. This room gives the children a pleasant place to play indoors as well.

Library

For most Dagstuhl Seminars, books available in our library that were written by seminar participants are displayed separately. On average, more than 50 books are presented per Dagstuhl Seminar.

In 2022, the comprehensive collection “Dagstuhl’s Impact”[8], which documents published results of Dagstuhl Seminars, grew substantially: 82 articles, 3 special journal issues, and 3 books were added.

For reasons of space, the library’s basement will be equipped with mobile shelving in the future. After an intensive planning phase, the preparatory work for this has begun. Following an inventory review, the entire book collection in the library basement was moved to other locations in the building, including all shelving, to make room for the new shelving system, which is scheduled to be installed in early 2023.

Art

Since 1995, Schloss Dagstuhl has hosted art exhibitions on a regular basis. Until now, most exhibitions were organized by individual artists or sometimes small groups of artists. The year under review saw the establishment of a cooperation between the institute of contemporary art in Saarland and Schloss Dagstuhl. The first fruit of this collaboration was the exhibition “THINKING STRUCTURES” with works by Monika von Boch, Jo Enzweiler, and Vera Molnár. We are thankful for the support of this collaboration provided by the Friends of Dagstuhl (Verein zur Förderung von Schloss Dagstuhl – Leibniz-Zentrum für Informatik e.V.).

Further information about the art program in general, the exhibition in 2022, as well as the newly established cooperation can be found in Chapter 10.

[8] http://www.dagstuhl.de/bibliothek/dagstuhls-impact/

2 Seminare und Workshops
Seminars and Workshops

2.1 Dagstuhl-Seminare

Die Dagstuhl-Seminare haben als wesentliches Instrument der Forschungsförderung Priorität bei der Gestaltung des Jahresprogramms. Hauptziel der Seminare ist die Unterstützung der Kommunikation und des Dialogs zwischen Wissenschaftlern, die an den Forschungsfronten von miteinander verknüpften Forschungsfeldern in der Informatik arbeiten. Die Seminare ermöglichen die Vorstellung neuer Ideen, die Diskussion von aktuellen Problemen sowie die Weichenstellung für zukünftige Entwicklungen. Sie bieten außerdem die Möglichkeit zum Austausch zwischen vielversprechenden Nachwuchswissenschaftlern und internationalen Spitzenforschern in einem speziellen Forschungsgebiet.

Die Teilnahme an den üblicherweise einwöchigen Seminaren ist nur auf persönliche Einladung durch Schloss Dagstuhl möglich. Das Zentrum übernimmt einen Teil der Kosten, sodass die besten Wissenschaftler einschließlich jungen Forschenden und Promovierenden teilnehmen können. Zu den ehemaligen Gästen zählen 27 Preisträger des Turing-Awards, der höchsten Auszeichnung, die im Bereich der Informatik auf internationaler Ebene verliehen wird.

Charakteristisch für Dagstuhl ist die Etablierung von richtungsweisenden sowie gebietsübergreifenden Seminaren. Manche Themen, die ausgiebig in Dagstuhl diskutiert wurden, entwickelten sich anschließend zu sehr aktiven Forschungsbereichen, die teilweise zu DFG-Schwerpunkten und anderen Förderprogrammen führten. Bei einer Reihe von Forschungsgebieten wurden durch Dagstuhl-Seminare Gruppen zusammengeführt, die zwar an verwandten Problemen und Verfahren forschen, denen aber bisher keine gemeinsame Diskussionsplattform zur Verfügung stand. Dies gilt insbesondere auch für Disziplinen, die nicht zur Informatik gehören. Wichtige Forschungsgebiete, für die in Dagstuhl bereits mehrfach eine intensive Zusammenarbeit mit der Informatik erschlossen und vertieft wurde, sind Biologie (seit 1992) und Sport (seit 2006). Die Themen der Dagstuhl-Seminare bieten eine hervorragende und sehr breite Übersicht über die aktuellen Forschungsgebiete der Informatik.

Jedes Dagstuhl Seminar wird gebeten, einen kurze Dokumentation zu erstellen, die eine Zusammenfassung des Seminarverlaufs, eine Kurzübersicht über die gehaltenen Vorträge und eine Zusammenfassung grundsätzlicher Ergebnisse enthält. Diese Berichte, die in der Zeitschrift *Dagstuhl Reports* veröffentlicht werden, gewährleisten eine hohe Sichtbarkeit und eine zeitnahe Kommunikation der Ergebnisse. Dagstuhl Reports wird jährlich in einem Band mit 12 Ausgaben veröffentlicht. Jede Ausgabe dokumentiert jeweils die Dagstuhl-Seminare und Dagstuhl-Perspektiven-Workshops eines Monats. Die Dagstuhl Reports sind über die Dagstuhl-Website frei zugänglich.[9]

Kapitel 6 enthält Zusammenfassungen der Dagstuhl-Seminare und Perspektiven-Workshops. Im Kapitel 14 sind alle Veranstaltungen, die 2022 stattfanden, aufgelistet. Auf der Dagstuhl-Website ist das Programm der kommen-

2.1 *Dagstuhl Seminars*

Dagstuhl Seminars, the center's key instrument for promoting research, are accorded top priority in its annual program. The central goal of the Dagstuhl Seminar program is to stimulate new research by fostering communication and dialogue between scientists working on the frontiers of knowledge in interconnected fields related to informatics. New ideas are showcased, topical problems are discussed, and the course is set for future development in the field. The seminars also provide a unique opportunity for the exchange of research views and findings in a specific cutting-edge field of informatics between promising young scientists and the international elite of the research area.

Participation in these events – which generally last one week – is possible only by way of personal invitation from Schloss Dagstuhl. The center assumes part of the associated costs in order to enable the world's most qualified scientists, including young researchers and doctoral students, to participate. Among Dagstuhl's past guests are 27 winners of the ACM Turing Award, the highest achievable award within the international computer science community.

Dagstuhl's distinguished accomplishment is to have established pioneering, interdisciplinary seminars that have virtually become institutions themselves. Many of the topics addressed in-depth at Dagstuhl have subsequently developed into highly active research fields, resulting in some cases in DFG priority programs and other grant and funding programs. Dagstuhl Seminars often succeed in bringing together scientists from a range of research areas and disciplines whose work overlaps with respect to issues, methods and/or techniques, but who have never previously entered into constructive dialogue with one another. This especially applies to disciplines outside of the field of informatics. Key research areas for which in-depth collaboration with informatics specialists was initiated and consolidated at Dagstuhl include biology (since 1992) and sports (since 2006). The spectrum of seminar topics provides an excellent and broad overview of the areas currently under discussion in the informatics arena.

Each Dagstuhl Seminar is asked to contribute a record of the seminar proceedings in the form of a Dagstuhl Report. The report gives an overview of the seminar's program, talks, and results in a journal-like manner to allow for a high visibility and timely communication of its outcome. The periodical *Dagstuhl Reports* is published in one volume with 12 issues per year; each issue documents the Dagstuhl Seminars and Dagstuhl Perspectives Workshops of a given month. Dagstuhl Reports are openly accessible and can be downloaded from the Dagstuhl website.[9]

Chapter 6 contains a collection of the summaries of the 2022 Seminars and Perspectives Workshops. Chapter 14 provides a comprehensive list of all events that took place during 2022, and a seminar program covering the upcoming 24 months is available on the Dagstuhl website.

den 24 Monate verfügbar.

2.2 Dagstuhl-Perspektiven-Workshops

In Ergänzung zu den Dagstuhl-Seminaren werden Dagstuhl-Perspektiven-Workshops veranstaltet, bei denen 25–30 ausgewiesene Forschenden ein bereits fest etabliertes Forschungsgebiet betreffende Tendenzen und neue Perspektiven der weiteren Entwicklung dieses Gebietes diskutieren. Im Gegensatz zu Dagstuhl-Seminaren werden statt aktueller Forschungsergebnisse im Wesentlichen Positionspapiere vorgetragen, welche den aktuellen Stand des Gebietes, offene Probleme, Defizite und vielversprechende Richtungen beschreiben. Der Fokus in den Workshops liegt auf Teilgebieten oder mehreren Gebieten der Informatik. Jeder Workshop hat zum Ziel

- den Stand eines Gebietes zu analysieren,
- Potenziale und Entwicklungsperspektiven bestehender Forschungsfelder zu erschließen,
- Defizite und problematische Entwicklungen aufzudecken,
- Forschungsrichtungen aufzuzeigen und
- Innovationsprozesse anzustoßen.

Die Dagstuhl-Perspektiven-Workshops, die 2022 statt fanden, sind in Fig. 2.1 aufgelistet.

Die Ergebnisse der intensiven Diskussionen werden in einem Manifest zusammengefasst, welches die offenen Probleme und die möglichen Forschungsperspektiven für die nächsten 5–10 Jahre aufzeigt. Dagstuhl koordiniert die gezielte Weitergabe dieses Manifests, um forschungsspezifische Impulse an deutsche und europäische Institutionen der Forschungsförderung zu geben (EU, BMBF, DFG, etc.). Kurzfassungen der Manifeste werden regelmäßig im Forum des *Informatik Spektrum* (Springer-Verlag) vorgestellt. Die vollständigen Manifeste werden in unserer Fachzeitschrift Dagstuhl Manifestos[10] veröffentlicht.

Eine Liste der vergangenen und kommenden Dagstuhl-Perspektiven-Workshops ist auf der Dagstuhl-Website verfügbar.[11]

2.2 *Dagstuhl Perspectives Workshops*

In addition to the traditional Dagstuhl Seminars, the center organizes Dagstuhl Perspectives Workshops. A Perspectives Workshop involves 25–30 internationally renowned senior scientists who wish to discuss strategic trends in a key research area that is already well established and to develop new perspectives for its future evolution. In contrast to Dagstuhl Seminars, Perspectives Workshops do not address current research results but reflect the overall state of a field, identifying strengths and weaknesses, determining promising new developments, and detecting emergent problems and synergies. The workshops tend to focus on subfields or are interdisciplinary in nature, thus covering more than one informatics field. Each workshop aims to

- contribute to an analysis of the present status of a field,
- tap into potentials and development perspectives of existing fields of research,
- detect shortcomings and problematic developments,
- show research directions, and
- trigger innovation processes.

Dagstuhl Perspectives Workshops held in 2022 are listed in Fig. 2.1.

The results of the in-depth discussions of each workshop are presented in a manifesto detailing open issues and possible research perspectives in that specific field for the coming 5–10 years. Schloss Dagstuhl coordinates the targeted dissemination of this manifesto as research policy impulses to German and other European research donors and sponsors (EU, German Federal Ministry of Education and Research, DFG, etc.). Short versions of the manifestos are regularly presented in a forum of the *Informatik Spektrum* journal (published by Springer); full versions of the manifestos are published in our periodical Dagstuhl Manifestos[10].

A list of past and upcoming Dagstuhl Perspectives Workshops can be found on our website.[11]

[9] https://www.dagstuhl.de/dagrep/
[10] https://www.dagstuhl.de/dagman
[11] https://www.dagstuhl.de/pw-list

Human-Centered Artificial Intelligence http://www.dagstuhl.de/22262
Current and Future Challenges in Knowledge Representation and Reasoning http://www.dagstuhl.de/22282

Fig. 2.1
Dagstuhl Perspectives Workshops held in 2022.

Einreichung der Anträge und Begutachtungsverfahren

2.3

Proposal Submission and Review Process

Die gleichbleibend hohe Qualität der Dagstuhl-Seminare und Dagstuhl-Perspektiven-Workshops wird durch Auswahl der Anträge gewährleistet, die aus Sicht von Schloss Dagstuhl das größte Potential haben, abseits etablierter Konferenzen neue und wichtige Forschungsprobleme mit Wissenschaftlern aus oft unterschiedlichen Gebieten zu identifizieren und zeitgleich mögliche Methoden und Lösungsansätze zu diskutieren.

Das Zentrum erbittet zweimal im Jahr Themenvorschläge von führenden Wissenschaftlerinnen und Wissenschaftlern aus der ganzen Welt, die ihre Seminaranträge zusammen mit einer vorläufigen Teilnehmerliste einreichen. Die Anträge werden dann vom Wissenschaftlichen Direktorium (siehe Kapitel 11.3) begutachtet und abschließend bei zweitägigen Sitzungen auf Schloss Dagstuhl intensiv diskutiert und über sie entschieden.

Es wird sicher gestellt, dass jedes Dagstuhl-Seminar durch ein starkes Organisatorenteam betreut wird, ein für die Informatik-Community relevantes Thema anspricht, ein kohärentes und gut strukturiertes wissenschaftliches Programm präsentiert und eine Gruppe von geeigneten Teilnehmerinnen und Teilnehmern zusammenbringt, deren kollektive Fachkenntnis einen bedeutenden Durchbruch in dem betreffenden Forschungsfeld ermöglichen kann. Zudem wird auf eine ausgeglichenen Repräsentation wissenschaftlicher Gemeinden, geographischer Regionen und besonders auf das Miteinbeziehen junger und weiblicher Wissenschaftler geachtet.

Die Informatikforscherinnen und Informatikforscher zeigten 2022 wieder ein hohes Interesse am Organisieren von Dagstuhl-Seminaren und Dagstuhl-Perspektiven-Workshops durch die Einreichung von insgesamt 93 Anträgen in den Antragsrunden im Januar und September 2022. Davon wurden 64 Anträge, was 69 % entspricht, angenommen. Die Vergleichsweise hohe Rate der angenommenen Anträge – in den vergangenen 7 Jahren variierte die Rate der angenommen Anträge zwischen 56 % und 69 % – ist auch darauf zurückzuführen, dass 11 wegen der Pandemie abgesagte oder nur stark eingeschränkt stattgefunden habene Seminare, die bereits einmal erfolgreich beantragt wurden, neu beantragt haben. (siehe Fig. 2.2).

Unter den neu genehmigten Dagstuhl-Seminaren und Dagstuhl-Perspektiven-Workshops gab es wie in den vergangenen Jahren wieder verschiedene Konstellationen bzgl. Dauer und Größe (vgl. Fig. 2.4). Von diesen waren 54 Seminare (hier und im Folgenden wird, sofern nicht anders angegeben, das Wort „Seminar“ sowohl für Dagstuhl-Seminare als auch für Dagstuhl-Perspektiven-Workshops verwendet) eigentlich für 2023 vorgesehen, doch nur 9 konnten tatsächlich für 2023 eingeplant werden. Eines konnte noch für 2022 eingeplant werden. Diese langen Vorlaufszeiten entstanden maßgeblich durch die pandemiebedingte Verschiebung und Absage von Seminaren in 2020 und 2021.

Schloss Dagstuhl maintains the high quality of the Dagstuhl Seminar and Dagstuhl Perspectives Workshop series by identifying those proposals that promise a high potential to engage researchers – often from different disciplines – in scientific discussions on new and important research problems and their most promising solutions, outside of the existing conferences.

The center solicits topics for new seminars and workshops twice a year from leading researchers worldwide, who submit their proposals together with a list of potential scientists to be invited. The proposals and suggested invitee lists are then reviewed by Dagstuhl's Scientific Directorate (see Section 11.3) and finally discussed and decided during a two-day meeting at Schloss Dagstuhl.

This process ensures that every Dagstuhl Seminar and Dagstuhl Perspectives Workshop is backed by a strong team of organizers, addresses a topic of relevance to the computer science community, presents a coherent and well-structured scientific agenda, and brings together the right group of participants whose collective expertise can lead to a significant breakthrough in the area to be addressed. The balance of research communities and geographical regions, and especially the inclusion of junior and female researchers, are also taken into account during the review process.

In 2022, computer science researchers again showed a great interest in organizing Dagstuhl Seminars and Dagstuhl Perspectives Workshops by submitting a total of 93 proposals in the January and September 2022 proposal rounds. Of these, 64 proposals were accepted, which corresponds to 69 %. This comparatively high rate of accepted proposals – in the past 7 years, the rate of accepted proposals varied from 56 % to 69 % – is also due to the fact that 11 seminars that had already been successfully applied for once, but were canceled or severely restricted due to the pandemic, have reapplied (see Fig. 2.2).

Among the newly accepted Dagstuhl Seminars and Dagstuhl Perspectives Workshops of 2022, there was – as in the previous years – a wide variation in terms of duration and size (see Fig. 2.4). Of these, 54 seminars (here and in the following, the word "seminar" is meant to include both Dagstuhl Seminars and Dagstuhl Perspektives Workshops, if not specified otherwise) were to take place in 2023; however, only 9 could be scheduled for 2023 and one could be scheduled for 2022. These long lead times were largely due to the number of postponed and canceled seminars from 2020 and 2021 due to the pandemic.

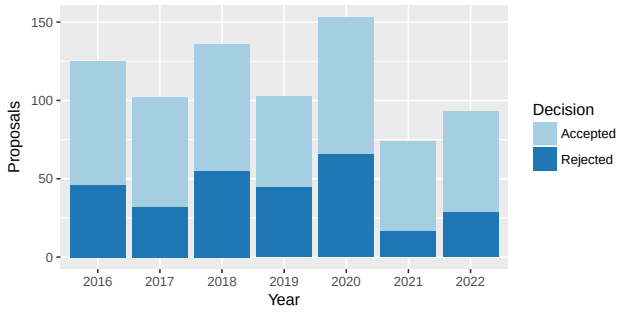


Fig. 2.2
Overview of proposed and accepted Dagstuhl Seminars and Dagstuhl Perspectives Workshops in 2016–2022.

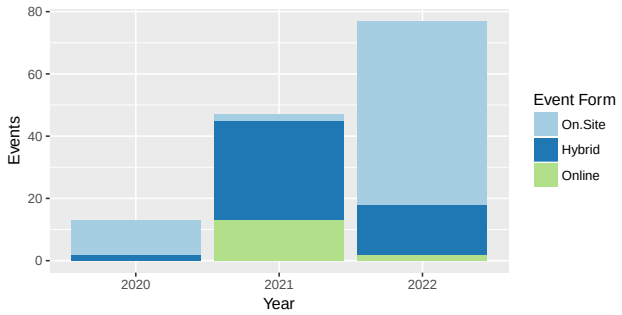


Fig. 2.3
Number of seminars in 2020–2022 by form.

Seminar-Programm 2022 — 2.4 — *The Seminar Program in 2022*

Hier und im Folgenden ist zu beachten, dass die Zahlen von 2022 durch die anhaltenden Auswirkungen der Pandemie immer noch nicht mit den Zahlen der Jahre vor 2020 vergleichbar sind.

Grundsätzlich kann Schloss Dagstuhl in jeder Woche zwei Seminare mit insgesamt etwa 75 Teilnehmenden beherbergen.

In 44 von 49 Wochen, in denen das Tagungszentrum 2022 geöffnet war, fand mindestens ein Dagstuhl-Seminar oder Dagstuhl-Perspektiven-Workshop statt. In 33 Wochen waren es sogar zwei. In 5 Wochen war das Zentrum nur durch andere Veranstaltungen belegt. Es blieb keine Woche unbelegt, in einer Woche war aber tatsächlich kein Gast vor Ort.

2022 fanden insgesamt 77 Dagstuhl-Seminare und Dagstuhl-Perspektiven-Workshops statt. In Fig. 2.6 ist die Entwicklung der vergangenen Jahre dargestellt.

Auch 2022 bot Schloss Dagstuhl allen Seminarorganisatoren die Option an, ihre Seminare hybrid – also mit einigen Teilnehmenden, die vor Ort sind und weiteren Teilnehmenden, die über Video-/Audioübertragung zugeschaltet sind. Auf Nachfrage wurde in begründeten Ausnahmefällen auch die Option reiner online Seminare angeboten. Fig. 2.3 gibt einen Überblick, wie viele Veranstaltungen rein vor Ort, hybrid oder rein online stattgefunden haben.

Here and in the following, it should be noted that the figures for 2022 are still not comparable with the figures for the years prior to 2020, due to the continuing effects of the pandemic.

In principle, Schloss Dagstuhl can host two seminars each week with a total of about 75 participants.

At least one Dagstuhl Seminar or Dagstuhl Perspectives Workshop was held in 44 of the 49 weeks for which the center was open in 2022. In 33 of those weeks, there were in fact two seminars in parallel. In 5 weeks, there were exclusively other types of events scheduled. There were no weeks without scheduled events, but in one week all events were purely online.

Altogether, there were 77 Dagstuhl Seminars and Dagstuhl Perspectives Workshops in 2022. Fig. 2.6 shows the development in recent years.

In 2022, Schloss Dagstuhl again offered all seminar organizers the option to hold a hybrid seminar – i.e., with some participants that are on site at Schloss Dagstuhl and some that are included using video and audio transmission. Upon request, the option for a purely online seminar was also offered in exceptional cases with justification. Fig. 2.3 provides an overview of how many events took place purely on site, in a hybrid format, or purely online.

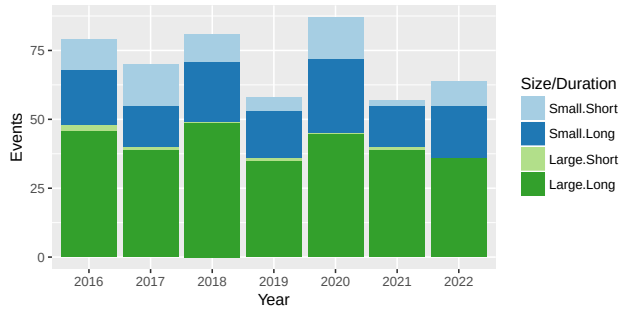


Fig. 2.4
Size and duration of Dagstuhl Seminars and Dagstuhl Perspectives Workshops accepted in 2016–2022.
Small = 30-person seminar, Large = 45-person seminar, Short = 3-day seminar, Long = 5-day seminar.

Angaben zu Teilnehmenden und Organisierenden

2.5

Participant and Organizer Data

Insgesamt nahmen 2811 Forschende an Veranstaltungen teil, davon 2278 an Dagstuhl-Seminaren und Perspektiven-Workshops. Immerhin waren 2386 Forschende – davon 1895 bei Dagstuhl-Seminaren und Perspektiven-Workshops – vor Ort. Damit haben nur 17 % der Teilnehmenden an Seminaren online teilgenommen.

Grundsätzlich werden hier alle Teilnehmende gezählt, unabhängig davon, ob sie vor Ort oder online teilgenommen haben. Einen Eindruck, wieviele Teilnehmende vor Ort waren bzw. remote teilgenommen haben, vermittelt Fig. 2.5.

Viele der internationalen Gäste der Seminare waren schon öfter in Dagstuhl. Dennoch zieht das Zentrum jedes Jahr auch neue Gesichter an, was den ständigen Wandel in der Forschung widerspiegelt. So nahmen, wie in den betrachteten Vorjahren, im Jahr 2022 wieder fast die Hälfte (49 %, 1 051 von 2 139) der Wissenschaftler und Wissenschaftlerinnen das erste Mal an einem Dagstuhl-Seminar oder Dagstuhl-Perspektiven-Workshop teil, während weitere 15 % an nur einem Seminar in den Jahren vorher teilgenommen hatten, weitere 9 % nur an zweien. Betrachtet man ausschließlich Teilnehmende, die vor Ort angereist waren, so erhält man ähnliche Ergebnisse, die nur um etwa 1 Prozentpunkt von den obigen abweichen.

Ein wenig andere Zahlen leiten sich aus unserer Gastumfrage ab. Hier ergibt sich, dass etwa 42 % der Antwortenden 2022 das erste Mal, 14 % zum zweiten Mal und weitere 9 % zum dritten Mal (siehe Fig. 2.7a) teilgenommen haben.

Ein beträchtlicher Anteil der Gäste besteht aus jungen Forschenden, die am Anfang ihrer Karriere stehen, und für die der Aufenthalt in Dagstuhl oftmals prägend ist für den weiteren Verlauf ihres Lebenswegs. Etwa 33 % der Gäste der Seminare im Jahr 2022, die an unserer Umfrage zur Qualitätskontrolle teilgenommen haben, stuften sich selbst als wissenschaftlichen Nachwuchs ein (siehe Fig. 2.7b). Wie in fast allen Vorjahren hatte Schloss Dagstuhl mit etwa einem Drittel eine ausgewogene Verteilung zwischen Nachwuchswissenschaftlern und erfahrenen Forschern, die im Laufe der Jahre (mit Ausnahme des letzten) relativ konstant geblieben war, was die Bemühungen des Zentrums zur Aufrechterhaltung der „Dagstuhl-Verbindung“ zwischen herausragenden jungen Wissenschaftlerinnen und Wissenschaftlern und ihren erfahrenen Kollegen zeigt.

Of the more than 2,278 scientists who participated in events, 2,811 participated in Dagstuhl Seminars and Dagstuhl Perspectives Workshops. 2,386 scientists participated on site, 1,895 of which were participating in Dagstuhl Seminars and Dagstuhl Perspectives Workshops. Thus, only 17 % of the seminar participants participated online.

In general, all participants are counted here, regardless of whether they participated on site or online. Fig. 2.5 illustrates how many participants were on site and how many participated remotely.

Participants in Dagstuhl Seminars come from all over the world, and a significant number of them choose to repeat the experience. Nevertheless, we see many fresh new faces every year, reflecting the changing informatics research across the globe. As in the previous years, in 2022, almost half (1,051 of 2,137, or 49 %) of the researchers were first-time visitors to Dagstuhl. About an additional 15 % of the participating researchers had already attended one previous seminar in the years before, and another 9 % had already attended two. Looking only at on-site participants, we get similar results, differing only in about 1 percentage point.

Slightly different numbers were obtained from our guest survey from 2022: about 42 % of the responders were first-time visitors, an additional 14 % were visiting for the second time, and yet another 9 % for the third time (see Figure 2.7a).

A substantial number of these guests were young researchers at the start of their careers, for whom the Dagstuhl experience can be of lifelong value. Approximately 33 % of the 2022 Dagstuhl Seminar survey respondents self-classified as junior (see Fig. 2.7b). As in almost all previous years (with the exception of last year), Schloss Dagstuhl had a balanced distribution between junior and senior researchers of about one-third, reflecting the center’s determined effort to maintain the “Dagstuhl connection” between brilliant junior scientists and their senior colleagues.

Mit 78 % war der Anteil von an Seminaren Teilnehmenden aus dem Ausland 2022 wieder auf einem gewohnt hohen Niveau. Das Diagramm in Fig. 2.7c zeigt die regionale Verteilung der Gäste für 2022 bei Seminaren. Mehr Details können Kapitel 13 entnommen werden.

In 2022 waren etwa 78 % aller Organisatorenteams des Seminar-Programms hinsichtlich des Geschlechts gemischt und rund 29 % aller Organisatoren waren Frauen (siehe Fig. 2.8a). Der Anteil an weiblichen Seminarteilnehmern war mit 28 % überdurchschnittlich hoch (siehe Fig. 2.8b).

At around 78 %, the proportion of guests with a non-German affiliation in Dagstuhl Seminars was back to its usual high level. The chart in Fig. 2.7c shows the regional distribution of our Dagstuhl Seminar guests in 2022. For a detailed breakdown, please refer to Chapter 13.

Im Jahr 2022, about 78 % of all organizer teams in our scientific seminar program were mixed with respect to gender and about 29 % of all organizers were women (see Fig. 2.8a). The proportion of female seminar participants was above average at 28 % (see Fig. 2.8b).

Themen und Forschungsgebiete 2.6

Die thematischen Schwerpunkte der Dagstuhl-Seminare und Dagstuhl-Perspektiven-Workshops werden von den internationalen Antragstellern identifiziert und dem wissenschaftlichen Direktorium zur Durchführung vorgeschlagen. Hierdurch wird die internationale Forschungsgemeinde aktiv in die Programmgestaltung eingebunden – zugleich ist gewährleistet, dass aufgrund der Expertise der Antragstellenden in ihren jeweiligen Forschungsgebieten immer brandaktuelle Themenschwerpunkte gesetzt werden.

Im Folgenden sind beispielhaft einige thematische Schwerpunkte und dazugehörige Seminare aufgeführt. Die Aufzählung der Themen und Seminare hat keinen Anspruch auf Vollständigkeit und ist lediglich ein Versuch, einen kurzen Einblick in das facettenreiche Seminar-Programm zu geben. Kapitel 6 bietet mit den Kurzzusammenfassungen der Seminare und Perspektiven-Workshops einen vollständigen Überblick über das wissenschaftliche Seminar-Programm des Jahres 2022.

Wie jedes Jahr deckten unsere Dagstuhl-Seminare große Teile der Informatik ab. Von den theoretischen Grundlagen, etwa *Finite and Algorithmic Model Theory* (22051) und *Algebraic and Analytic Methods in Computational Complexity* (22371) über mehr angewandte Themen wie *Computational Models of Human-Automated Vehicle Interaction* (22102) und *Framing in Communication: From Theories to Computation* (22131) bis zu praktisch relevanten Themen wie *Database Indexing and Query Processing* (22111) und *The Human Factors Impact of Programming Error Messages* (22052) war wieder alles vertreten. Anwendungen vom omnipräsenten Thema Machine Learning waren auch Bestandteil einiger Seminare wie zum Beispiel *Developmental Machine Learning: From Human Learning to Machines and Back* (22421). Auch gesellschaftspolitische relevante Fragen werden in Dagstuhl diskutiert, wie die Seminare *AI for the Social Good* (22091) und *Technologies to Support Critical Thinking in an Age of Misinformation* (22172), zeigen.

Diese kleine Auswahl von Seminaren soll aber nicht darüber hinwegtäuschen, dass jedes der im Jahr 2022 veranstalteten Seminare wichtige Themen adressiert hat, die von den beteiligten Wissenschaftlern mit großem Engagement diskutiert wurden und so die weitere Entwicklung in den einzelnen Gebieten wieder ein gutes Stück weitergebracht hat.

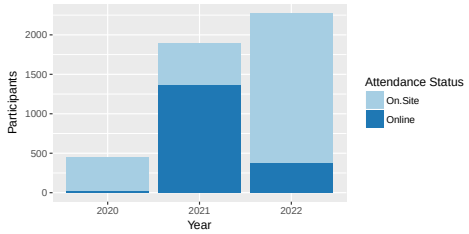
Topics and Research Areas

The topics of Dagstuhl Seminars and Dagstuhl Perspectives Workshops are identified by researchers from all over the world, who pass on this information to the Schloss Dagstuhl Scientific Directorate in their submitted proposals. The international research community is thus actively involved in shaping Dagstuhl's scientific seminar program, and their expertise ensures that the most important cutting edge topics are emphasized.

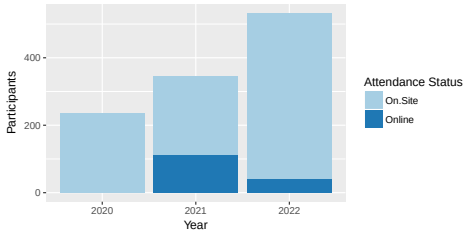
The following overview gives some topical focal points and a few respective seminars from 2022. Neither the list of focal points nor the list of seminars is exhaustive. It merely attempts to offer a brief insight into the multifaceted seminar program of 2022. Chapter 6 provides a complete overview of the 2022 scientific seminar program including short summaries of the Seminars and Perspectives Workshops.

As in every year, the topics of the Dagstuhl Seminars covered large parts of computer science. Everything was represented again, from the theoretical basics, such as *Finite and Algorithmic Model Theory* (22051) and *Algebraic and Analytic Methods in Computational Complexity* (22371), to more applied topics like *Computational Models of Human-Automated Vehicle Interaction* (22102) and *Framing in Communication: From Theories to Computation* (22131), to practically relevant topics like *Database Indexing and Query Processing* (22111) and *The Human Factors Impact of Programming Error Messages* (22052). Applications of the omnipresent topic of machine learning were also a part of some seminars like for example *Developmental Machine Learning: From Human Learning to Machines and Back* (22421). Questions of societal relevance are being discussed in Dagstuhl as well, as demonstrated by the seminars *AI for the Social Good* (22091) and *Technologies to Support Critical Thinking in an Age of Misinformation* (22172).

This brief selection of seminars should not draw attention from the fact that each of 2022's seminars addressed important topics which were discussed by the involved researchers with great commitment and hence pushed forward the development in the individual areas.



(a) Distribution of participants of seminars in group A.



(b) Distribution of participants of events in group B.

Fig. 2.5

Number of participants by attendance status and group. Group A = Dagstuhl Seminars and Dagstuhl-Perspectives-Workshops. Group B = all other events (GI-Dagstuhl Seminars, educational events, and research group meetings).

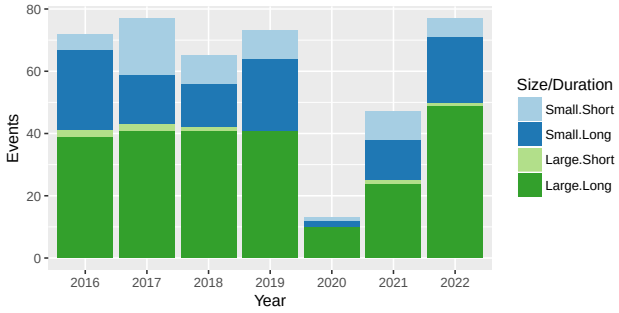
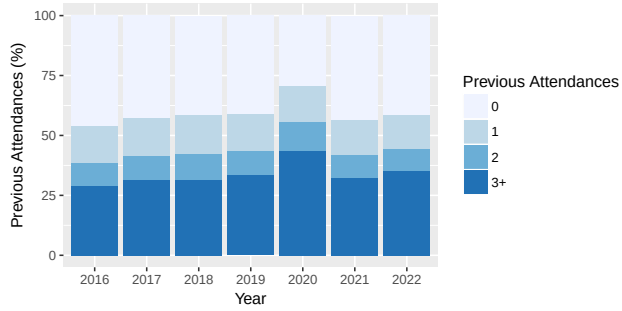
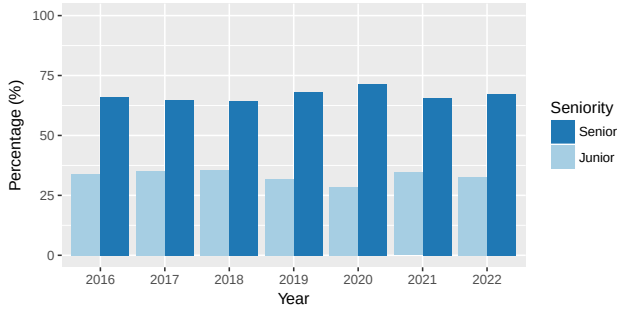


Fig. 2.6

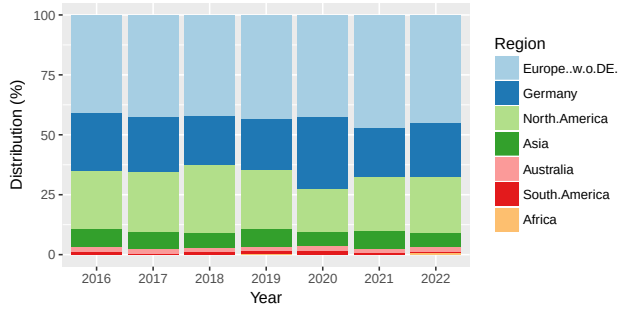
Size and duration of Dagstuhl Seminars and Dagstuhl Perspectives Workshops held in 2016–2022.
Small = 30-person seminar, large = 45-person seminar, short = 3-day seminar, long = 5-day seminar.



(a) Distribution of the number of previous attendances of participants, according to survey data.



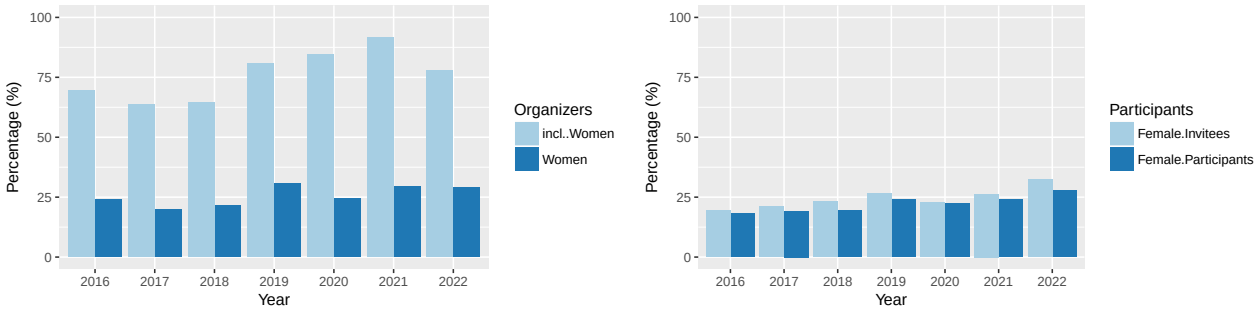
(b) Percentage of junior researchers, according to survey data.



(c) Distribution of the origin of participants by region.

Fig. 2.7

Participants of Dagstuhl Seminars and Dagstuhl Perspectives Workshops in 2016–2022.



(a) Female organizers and organizer teams including women.

(b) Female invitees and participants.

Fig. 2.8

Female researchers at Dagstuhl Seminars and Dagstuhl Perspectives Workshops in 2016–2022.

Weitere Veranstaltungstypen 2.7

Neben den Dagstuhl-Seminaren und Dagstuhl-Perspektiven-Workshops finden noch weitere Veranstaltungen im Zentrum statt. Zu diesen Veranstaltungen gehören:

- GI-Dagstuhl-Seminare, die den wissenschaftlichen Nachwuchs zu einem bestimmten Thema zusammenführen. Sie werden in Kooperation mit der GI durchgeführt und von dieser sowie von Dagstuhl gefördert. Anträge auf GI-Dagstuhl Seminare werden vom Vorstand der GIBU (GI Beirat der Universitätsprofessoren) und vom Wissenschaftlichen Direktor von Schloss Dagstuhl begutachtet.
- Weiterbildungsveranstaltungen wie Sommerschulen und Lehrerfortbildungen.
- Forschungsgruppentreffen wie Klausurtagungen von Graduiertenkollegs, GI-Fachgruppen und anderen akademischen Arbeitsgruppen.
- Forschungsaufenthalte von Einzelpersonen, die sich für eine oder mehrere Wochen für intensive Studien nach Dagstuhl in Klausur zurückziehen.

Further Event Types

In addition to Dagstuhl Seminars and Dagstuhl Perspectives Workshops, Schloss Dagstuhl hosts a number of further events, including:

- GI-Dagstuhl Seminars bring young scholars together to discuss and learn about a specific topic. They are run and sponsored by the German Informatics Society (GI) in association with Schloss Dagstuhl. Proposals for GI-Dagstuhl Seminars are reviewed by the managing board of the GIBU (GI advisory board of computer science professors) and the Scientific Director of Schloss Dagstuhl.
- continuing education courses including summer schools and vocational training for teachers.
- research group meetings including conferences of graduate research training groups, GI specialist groups, and other academic working groups.
- research stays of scientists who wish to use the center as a retreat for several weeks in order to devote themselves to their studies undisturbed.

Qualitätssicherung 2.8

Schloss Dagstuhl befragt die Teilnehmenden der Dagstuhl-Seminare und der Dagstuhl-Perspektiven-Workshops mit Hilfe eines Fragebogens zu ihrer Zufriedenheit mit inhaltlichen und organisatorischen Aspekten ihres Dagstuhlbesuchs. Die Ergebnisse jedes Fragebogens werden im Haus wöchentlich allen Abteilungen zugänglich gemacht, um eine schnelle Reaktion auf Probleme und Wünsche zu erreichen. Gleichzeitig werden anonymisierte Ergebnisse von inhaltlichen Fragen den Teilnehmern eines Seminars per E-Mail mitgeteilt, typischerweise in der Woche nach ihrem Aufenthalt. So erhalten alle Rückmeldungen über den Verlauf des Seminars und Hinweise für die Organisation von zukünftigen Seminaren. In den zur Verfügung gestellten PDF-Dokumente werden die statistischen Ergebnisse mit Hilfe von aussagekräftigen Diagrammen aufbereitet.

Fig. 2.9 zeigt die Zufriedenheit der Antwortenden im Jahr 2022 zu ausgewählten Aspekten ihres Aufenthaltes. Grundlage ist die Auswertung von 1 153 Fragebögen,

Quality Assurance

The center conducts surveys of the participants of the Dagstuhl Seminars and Dagstuhl Perspectives Workshops, the questionnaire containing questions about their satisfaction with the content of the event and the organization of their visit. The results of each questionnaire are made available to all of the center's departments every week, thus enabling a quick response to issues and requests. At the same time, anonymized results of the content questions are made available to the seminar participants via e-mail, typically in the week following their stay at the center. This enables the organizers to receive feedback on how the seminar went and tips for organizing future seminars. In the pdf files with the results, the statistics are visualized using illuminative diagrams.

Fig. 2.9 shows the satisfaction of responding participants in 2022 with regard to selected aspects of their stay. The results were compiled from 1,153 questionnaires, representing the responses of about 51 % of all 2,278 participants. These excellent results are not only a recognition of

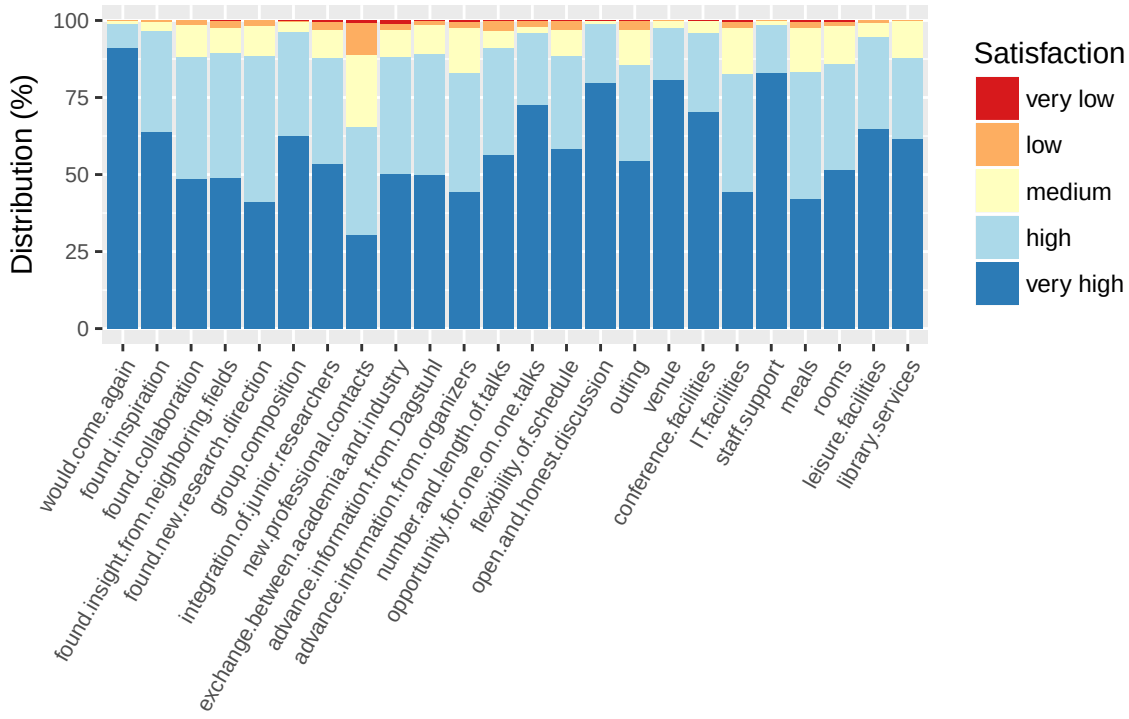


Fig. 2.9
Satisfaction of Dagstuhl Seminar and Dagstuhl Perspectives Workshop participants in 2022. According to survey results.

welche die Meinung von etwa 51 % der 2 278 Teilnehmenden repräsentieren. Das durchweg sehr gute Ergebnis ist Anerkennung und Herausforderung zugleich. Die Rücklaufquote von etwa 50 % war in 2022 wieder gewohnt hoch, während sie 2021 mit etwa 35 % ungewohnt niedrig war.

Als Teil des Einladungsprozesses der Dagstuhl-Seminare und der Dagstuhl-Perspektiven-Workshops wird die Liste der von dem Organisatorenteam zur Einladung Vorgeschlagenen von Schloss Dagstuhl auf eine ausgewogene Zusammensetzung geprüft, bevor Schloss Dagstuhl Einladungen ausspricht. Mittels einer täglich aktualisierten Webseite bietet Schloss Dagstuhl allen Organisatorenteams einen direkten Einblick in den Status der eingeladenen Gäste bezüglich Zu- oder Absage.

the center's past work but also pose a challenge to its future work. The response rate of about 50 % was again high in 2022, while it was unusally low in 2021 with about 35 %.

During the invitation process for Dagstuhl Seminars and Dagstuhl Perspectives Workshops, the Organizers compile a list of proposed invitees which is reviewed by Schloss Dagstuhl to check it for a balanced composition before Schloss Dagstuhl extends invitations. Via a dedicated webpage that is updated daily, Schloss Dagstuhl gives the organizers direct access to view the status of invitee replies.

Auslastung des Zentrums 2.9 *Utilization of the Center*

Im Jahr 2022 war die Auslastung deutlich höher als in den beiden vergangenen Jahren, die ganz im Zeichen der Covid-19-Pandemie und Reisebeschränlungen standen, erreichte aber bei weitem noch nicht das Niveau von 2019, dem letztem Jahr vor der Pandemie. Es gab 2022 insgesamt 10 009 Gasttage, wobei 8 687 Gasttage auf Dagstuhl-Seminare und Dagstuhl-Perspektiven-Workshops entfielen. Im Gegensatz zu den meisten anderen Statistiken zählen hier nur die Tage, an denen die Gäste tatsächlich vor Ort waren. Im Bereich Seminare war damit die Anzahl der Gasttage im Vergleich zu 2019 um etwa 1/4 reduziert.

In 2022, occupancy was significantly higher than in the previous two years, which were dominated by the Covid-19 pandemic and travel restrictions. But it still fell far short of the levels seen in 2019, the last year before the pandemic. There were 10,009 overnight stays in total in 2022, with 8,687 overnight stays for a Dagstuhl Seminar or a Dagstuhl Perspectives Workshop. In contrast to most of the other statistics, only days in which the guests where actually on site are counted here. In the seminars segment, the number of overnight stays was thus reduced by about 1/4 compared to 2019.

Es fanden im Berichtsjahr insgesamt 105 Veranstaltungen mit insgesamt 450 Veranstaltungstagen und insgesamt 2386 Gästen vor Ort statt. Weitere 425 Gäste von 38 Veranstaltungen nahmen online teil. Weitere Details können Kapitel 13 entnommen werden.

Die Wochenenden blieben 2022 ebenso unbelegt wie eine Wochen zu Weihnachten und 2 Wochen im Sommer, welche zu Instandhaltungs- und Verwaltungsarbeiten benötigt wurden. In 48 der 49 zur Verfügung stehenden Wochen waren tatsächlich Gäste vor Ort. In einer Woche gab es nur ein rein online stattfindendes Seminar.

Ein umfassendes Verzeichnis aller Veranstaltungen auf Schloss Dagstuhl im Jahr 2022 einschließlich Dagstuhl-Seminaren, Dagstuhl-Perspektiven-Workshops, GI-Dagstuhl-Seminaren und Veranstaltungen (z.B. Sommerschulen), bei denen Schloss Dagstuhl nur Veranstaltungsort war, findet sich in Kapitel 14. Auf unserer Webseite ist ein Kalender[12] verfügbar, in welchem die anstehenden Veranstaltungen eingesehen werden können, ebenso wie weitere Informationen und Materialien zu allen vergangenen, aktuellen und zukünftigen Veranstaltungen.

The center hosted a total of 105 events over a total of 450 days. In total, there were 2,386 guests on site. A further 425 guests, distributed over 38 events, participated online. See Chapter 13 for further details.

Weekends were kept free in 2022, as well as two weeks in the summer, which were needed for maintenance work to building facilities and administrative work, and one week at Christmas time. In 48 of the 49 available weeks there were actually guests on site. In one week, there was only a purely online seminar.

A comprehensive listing of all events at Schloss Dagstuhl in 2022, including Dagstuhl Seminars, Dagstuhl Perspectives Workshops, GI-Dagstuhl Seminars, and host-only events such as meetings and summer schools can be found in Chapter 14. See the Schloss Dagstuhl website to view our calendar[12] of upcoming events and further information and materials on all events: past, present, and future.

[12] https://www.dagstuhl.de/seminars/seminar-calendar/

3 Bibliographiedatenbank dblp
dblp computer science bibliography

3.1 Offene Bibliographiedaten für die Informatik

Moderne Informatik-Forschung benötigt den unmittelbaren und umfassenden Zugriff auf aktuelle Publikationen, um den Bedürfnissen in einer immer komplexer werdenden Forschungslandschaft gerecht zu werden. Auch bei der Einschätzung von Forschungsleistung ist die Verfügbarkeit verlässlicher Publikationsdaten unverzichtbar. Hoch qualitative und vollständige Metadaten sind in der Regel jedoch nur sehr schwer zu erhalten. Freie Suchmaschinen wie etwa Google erlauben einen weiten Einblick in das Internet, besitzen aber keinerlei Qualitätsgarantien oder semantische Organisation. Kommerzielle Datenbanken verkaufen Metadaten als teure Dienstleistung, weisen aber in vielen Fachdisziplinen (wie etwa in der Informatik) nur eine mangelhafte Abdeckung und eine oft ungenügende Datenqualität auf. Insbesondere die einzigartige Publikationskultur der Informatik mit ihrem Schwerpunkt auf Konferenzpublikationen bleibt dabei unberücksichtigt. Universitäten und außeruniversitäre Forschungseinrichtungen bemühen sich oftmals mit immensem personellen und finanziellen Aufwand und unter Belastung der einzelnen Forschenden, eigene Daten zu erheben. Diese Datensätze weisen jedoch zwangsläufig einen lokalen Einschlag auf und vermögen es nicht, ein detailliertes Bild einer Forschungsdisziplin als Ganzes zu zeichnen.

Die *dblp computer science bibliography* leistet auf diesem Gebiet nun bereits seit über 25 Jahren einen substanziellen Beitrag durch die offene Bereitstellung qualitätsgeprüfter und aufbereiteter Publikationsdaten für die gesamte Informatik. Dabei unterstützt dblp die Informatik-Forschung auf gleich mehreren Ebenen, etwa durch:

- Unterstützung der täglichen Forschungsarbeit, etwa bei der Literaturrecherche und dem Bezug von verfügbaren Volltexten
- Unterstützung des wissenschaftlichen Publikationsprozesses durch die Bereitstellung normierter bibliographischer Referenzdaten
- Unterstützung von Forschenden und Institutionen bei der Berichtspflicht durch die Sammlung und Aufbereitung von qualitätsgesicherten Publikationslisten
- Unterstützung von Forschungsfördernden und Entscheidungstragenden durch das öffentliche Verfügbarmachen von nach Daten-Facetten aufgeschlüsselten Publikationsnachweisen
- Ermöglichen einer Betrachtung der Informatik-Disziplin als Ganzes durch die Bereitstellung semantisch reichhaltiger Daten in Forschungs-Wissensgraphen

Darüber hinaus ist der dblp-Datensatz selbst Untersuchungsgegenstand mehrerer tausend Fachartikel.[13] Insgesamt ist dblp daher für die Informatik sowohl als Recherche-Tool, aber auch als Forschungsdatensatz, unverzichtbar geworden.

3.1 *Open Bibliographic Data in Computer Science*

Modern computer science research requires immediate and comprehensive access to current publications to meet the needs of an ever faster evolving and ever more complex research landscape. Not only in the everyday work of a researcher but also in the assessment of research performance, the availability of reliable bibliographic metadata has become indispensable. However, high-quality and complete metadata is very difficult to obtain. Free search engines like Google allow a broad insight into the Internet but have neither guarantees of quality nor any semantic organization. Commercial databases sell metadata as an expensive service, but in many disciplines (such as in computer science) their coverage is insufficient and the data quality is quite poor. In particular, the unique publication culture of computer science with its emphasis on conference publications remains disregarded. Most universities and non-university research institutions endeavor to collect their own data, yet often consume enormous human and financial resources and impose a burden on the individual researchers. However, these local data sets do inevitably have a local bias and are not suited to draw a detailed picture of a research discipline as a whole.

For over 25 years now, the *dblp computer science bibliography* has substantially contributed to solving this dilemma in the field of computer science by providing open, quality-checked, and curated bibliographic metadata. The dblp web service supports the computer science research community on several levels, for example by:

- supporting researchers in their daily work, e.g., when reviewing the literature or searching for full-text research articles
- supporting the scientific publication process by providing standardized bibliographic reference data
- supporting researchers and institutions in their reporting duties by collecting and editing quality-assured bibliographies
- supporting research funders and decision-makers, e.g., by providing publicly available and explorable bibliographic references
- enabling the study of the computer science discipline as a whole; e.g. by providing semantically meaningful data within a research knowledge graph

In addition, the dblp data set itself is object of study of several thousand research articles.[13] Hence, dblp has become indispensable to the computer science community as both a research tool and a research data set.

[13] Google Scholar liefert zum Suchbegriff „dblp“ über 54 000 Treffer, Semantic Scholar liefert 2 620; im Einzelnen weisen SpringerLink ca. 6 200 Artikel, Elsevier ScienceDirect über 1 300 Artikel, die ACM Digital Library ca. 5 600 Artikel und IEEE Xplore über 3 400 Artikel nach.
The search term „dblp“ results in more than 54 000 hits at Google Scholar and 2 620 hits at Semantic Scholar; in particular, SpringerLink lists about 6 200 articles, Elsevier ScienceDirect lists more than 1 300 articles, the ACM Digital Library lists 5 600 articles, and IEEE Xplore lists more than 3 400 articles.

Schloss Dagstuhl und dblp

3.2 Schloss Dagstuhl and dblp

Bereits seit Ende 2010 engagiert sich Schloss Dagstuhl für die ursprünglich an der Universität Trier entwickelte Bibliographiedatenbank dblp. Zunächst durch ein Projekt im Leibniz-Wettbewerb gefördert, wurde die Datenbank seit Juni 2013 von Schloss Dagstuhl direkt mitfinanziert. Im Zuge der Konsolidierung der Zusammenarbeit mit der Universität Trier wurden unter dem Dach von Schloss Dagstuhl dauerhafte Personalstellen im wissenschaftlichen Stab geschaffen, die hauptamtlich mit der Betreuung und Weiterentwicklung von dblp beauftragt sind. Ein eigens gegründeter dblp-Beirat (siehe Fig. 3.1) leistet seit 2011 die wissenschaftliche Aufsicht und unterstützt das dblp-Team mit seiner Expertise.

Pünktlich zum 25-jährigen Jubiläum von dblp erfolgte im November 2018 die endgültige Staffelübergabe des Betriebes der Datenbank von der Universität Trier an Schloss Dagstuhl. Damit einhergehend wurden von Bund und Ländern weitere Mittel für den Betrieb von dblp bereitgestellt und eine neue, eigens eingerichtete Außenstelle von Schloss Dagstuhl auf dem Campus II der Universität Trier angesiedelt. Betrieb und Erforschung der Datenbank erfolgen dabei weiterhin in enger Kooperation mit dem Fach Informatikwissenschaften der Universität Trier.

Das dblp-Team besteht mitlerweile aus 9,5 Vollzeitäquivalenten, welche an der redaktionellen, technischen und wissenschaftlichen Verbesserung der Infrastruktur arbeiten. Das Team konnte 2022 erneut eine Rekordzahl an mehr als 530 000 neue Publikationen indexieren (siehe Abschnitt 3.3). Gleichzeitig wurde die hohe Anzahl an behandelten Fehlerkorrekturfällen im Bestand gehalten (siehe Abschnitt 3.4). Parallel hierzu wird kontinuierlich an neuen Funktionen gearbeitet.

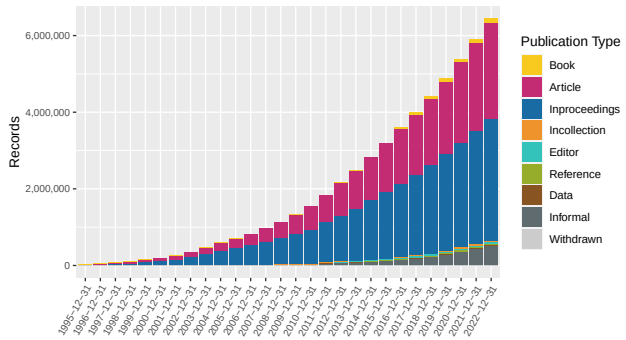
The cooperation between Schloss Dagstuhl and the dblp computer science bibliography – originally developed at the University of Trier – has been existing since late 2010. The commitment of Schloss Dagstuhl to dblp, initially funded by a project of the Leibniz Competition, is beeing funded directly by Schloss Dagstuhl since June 2013. As part of the consolidation of this cooperation, permanent scientific staff positions – assigned fulltime to the support and development of dblp – were created. The dblp advisory board (cf. Figure 3.1), established in 2011 at Schloss Dagstuhl, provides scientific supervision and supports dblp with its expertise.

In November 2018, the transfer of the database from the University of Trier to the Leibniz Center for Informatics at Schloss Dagstuhl took place just in time for dblp's 25th anniversary. At the same time, Dagstuhl's funding was increased to support the operation of dblp and a new Schloss Dagstuhl branch office for the dblp team was established on Campus II of the University of Trier. The database will continue to be operated and researched in close cooperation with the University's Department of Computer Sciences.

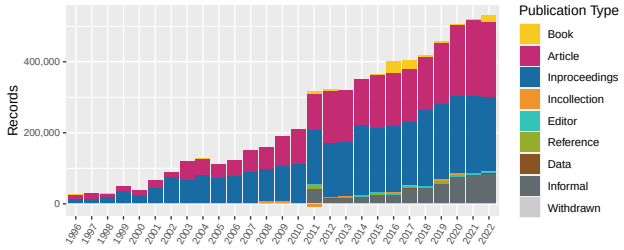
The dblp team – which had been a one-person project for more than a decade – now consists of 9.5 full-time equivalent staff members working on the editorial, technical, and scientific improvement of the infrastructure. In 2022, the team handled the indexing of more than 530,000 new publications (see Section 3.3) while at the same time kept handling a high rate of curation cases in the existing data (see Section 3.4). Parallel to the ongoing work with the data, services are continuously improved and new features are implemented.

dblp-Beirat | dblp Advisory Board
Prof. Dr. Hannah Bast University of Freiburg, Germany | *Chair*
Prof. Dr. Guillaume Cabanac Paul Sabatier University, Toulouse, France
Dr. Martin Fenner Front Matter, Münster, Germany
Prof. Dr. Silvio Peroni University of Bologna, Italy
Lydia Pintscher Wikimedia Deutschland – Association for the Promotion of Free Knowledge e.V., Berlin, Germany
Prof. Dr. Ruzica Piskac Yale University, New Haven, United States of America
Prof. Dr. Rüdiger Reischuk University of Lübeck, Germany
Prof. Dr.-Ing. Ralf Schenkel University of Trier, Germany
Prof. Raimund Seidel, Ph.D. Saarland University, Saarbrücken, Germany

Fig. 3.1
The dblp Advisory Board in 2022.



(a) Total number of records by year and type



(b) New records by year and type

Fig. 3.2
Development of the dblp data corpus.

Statistiken der Datenakquise 3.3 *Data Acquisition Statistics*

Die Bibliographiedatenbank dblp indexiert Publikationen anhand vollständiger Inhaltsverzeichnisse von Konferenzbänden oder Journalausgaben. Mit Hilfe einer eigens entwickelten Software zur Datenextraktion werden Metadaten von Verlagswebseiten ausgelesen und zur weiteren Bearbeitung vorbereitet. Die Metadaten werden anschließend vom dblp-Team redaktionell bearbeitet: Eventuelle Fehler werden korrigiert, mehrdeutige und ungenaue Angaben werden verbessert. Diese Datenpflege wird zwar von Hilfssoftware unterstützt, erfolgt aber vornehmlich händisch durch die jeweiligen Mitarbeitenden.

Dabei erwies sich 2022 erneut als das produktivste Jahr in der Geschichte der Datenbank. So wurden innerhalb von zwölf Monaten mehr als 530 000 neue Publikationen indexiert. Das entspricht mehr als 2 100 neuen Publikationen pro Arbeitstag. Die Neuaufnahmequote übertrifft damit zum sechsten Mal in Folge die Rekordzahl aus dem vorausgangenen Jahr. Die neu aufgenommenen Einträge verteilen sich zu 39,3% auf Konferenzbeiträge, zu 39,9% auf Journalartikel, zu 15,6% auf Preprints und „graue“ Literatur, zu 3,8% auf Monographien und Dissertationen, sowie zu 1,5% auf andere Publikationstypen. Am Ende des Jahres waren über 6,44 Millionen Publikationen aus den verschiedenen Teilgebieten der Informatik indexiert.

Ein Überblick über die Entwicklung der Datenakquise kann Fig. 3.2a und Fig. 3.2b entnommen werden.

The dblp computer science bibliography indexes conferences and journals on a per-volume basis. Using dblp's own web harvesting software, bibliographic metadata of journal or proceedings volumes are extracted from the publisher's website. This metadata is diligently checked and corrected by the dblp team. The data-cleaning process is assisted by algorithms but executed almost exclusively by hand.

The past year once again proved to be the most productive year in the history of dblp. Within 12 months, more than 530,000 new publications were indexed. This figure corresponds to more than 2,100 new records for each working day of the year. Hence, for the sixth year in a row, the rate of new additions has exceeded the record number from the previous year. This year's new records consist of 39.3% conference papers, 39.9% journal articles, 15.6% preprints and "grey" literature, 3.8% monographs and PhD theses, and 1.5% further publication types. By the end of 2022, a total of more than 6.44 million publications from all subdisciplines of computer science were indexed by dblp.

The development of the dblp data set is summarized in Figure 3.2a and Figure 3.2b.

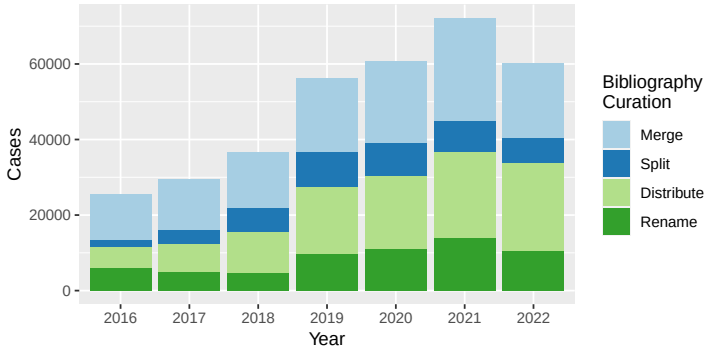


Fig. 3.3

Curation of existing dblp author bibliographies. The figures give the number of distinct edit cases (measured between the first and the last day of every given year) where a dblp team member manually corrected the assignment of publications within dblp author bibliographies. We distinguish between four curation cases: *Merge* = Two or more synonymous bibliographies have been merged into a single bibliography. *Split* = A single, homonymous bibliography has been split into two or more bibliographies. *Distribute* = A mixed case where records from two or more bibliographies have been redistributed between two or more bibliographies. *Rename* = A case where no actual publications have been reassigned, but the surface form of the author name(s) of a bibliography have been corrected or improved.

Statistiken der Datenkuration 3.4 *Data Curation Statistics*

Ein Hauptziel der intensiven Datenpflege des dblp-Teams ist es sicherzustellen, dass die bereitgestellten Bibliographien so korrekt und vollständig wie möglich sind. Das bedeutet, dass alle Publikationen eines Forschenden in nur einer einzigen Bibliographie aufgeführt sein sollen und dass diese Bibliographie auch nur Publikationen des jeweiligen Forschenden listen soll. Dies zu gewährleisten kann ziemlich herausfordernd sein, und trotz der Bemühungen des Teams werden regelmäßig Publikationen einer falschen Bibliographie zugeordnet. Aus diesem Grund wird der dblp-Datenbestand kontinuierlich von der dblp-Redaktion überprüft und eventuelle Zuordnungs- oder Daten-Fehler korrigiert.

Während spezielle Algorithmen dem Team helfen, solche Defekte aufzudecken, werden Korrekturen immer auf der Grundlage der Entscheidung eines menschlichen Kuratierenden durchgeführt. Dies ist notwendig, da die verfügbaren Metadaten in der Regel nicht genügend Informationen enthalten, um eine hochpräzise automatisierte Lösung zu erlauben. Oft ist eine manuelle Recherche unter Berücksichtigung externer Ressourcen erforderlich.

Im Laufe des Jahres 2022 wurden erneut mehr als 60 000 Fehlerfälle bearbeitet. Fig. 3.3 zeigt die Anzahl der gelösten Fehlerfälle in den letzten Jahren. Hauptfaktoren für die kontinuerliche Steigerung sind verbesserte interne Werkzeuge zum Erkennen potentiellen Fehlern und eine verbesserte Verfügbarkeit von ORCID[14]-Daten. ORCIDs sind eindeutige Kennungen, die Forschende mit ihren Publikationen verbinden können. Durch den Vergleich von ORCID-Daten mit dblp-Bibliographien gelingt es, in großer Zahl Zuordnungsfehler aufzudecken.

One main goal of the intensive data curation at dblp is to ensure that our author bibliographies are as correct and complete as possible. This means that all publications of a person should be listed in a single bibliography, and that a bibliography should only list publications from that specific author. It can be quite challanging to ensure this level of quality, and despite the best efforts of the dblp team, we regularly assign publications to the wrong bibliography. Because of this, our editorial team constantly checks our data and corrects such defects.

While specialized algorithms help the team to uncover and identify the nature of defects in our data, corrections are always executed based on the decision made by a human curator. This is necessary since the available metadata usually does not carry enough information to allow for a highly precise automated solution, and often requires a manual investigation taking external resources into account.

In 2022, more than 60,000 error cases have been handled and resolved by the dblp team. See Fig. 3.3 for a chart of the recnt years. Here, the team makes use of efficient internal tools that point out potential errors in the dblp corpus. On major component that enables theese tools is the increased availability of ORCID[14] data. ORCID is a unique contributor identifier that authors can attach to their publications. Comparing ORCID data with our bibliographies significantly helps us to identify errors in large numbers.

[14] https://orcid.org

	Trier		Dagstuhl		Total		
	2021	2022	2021	2022	2021	2022	%
page views per day	644,837	627,039	672,274	866,072	1,317,111	1,493,111	+13.4
user sessions (visits) per day	20,134	36,639	31,537	49,453	51,672	86,093	+66.6
page views per user session	32.0	17.1	21.3	17.5	25.5	17.3	-32.0
distinct users (IPs) per month	257,769	335,436	492,268	517,785	750,038	853,222	+13.8
data served per month	2,073.9 GB	1,793.7 GB	1,433.5 GB	1,378.1 GB	3,507.4 GB	3,171.7 GB	-9.6

Fig. 3.4
Average usage of the public dblp web servers. Trier = dblp.uni-trier.de, Dagstuhl = dblp.dagstuhl.de.

Nutzungsstatistiken 3.5 *Usage Statistics*

Im Jahr 2022 wurden vom dblp-Team zwei offizielle dblp-Server geführt. Die Informationen auf diesen Servern werden täglich aktualisiert und miteinander synchronisiert:

- Server Trier: dblp.uni-trier.de
- Server Dagstuhl: dblp.dagstuhl.de

Die allgemeine Adresse dblp.org ist dabei ein Alias für den dblp-Server in Dagstuhl.

Dabei war Server Trier in der Vergangenheit aufgrund seiner prominenten Sichtbarkeit in den Google-Suchergebnissen die mit Abstand bekannteste Adresse. Im Laufe des Jahres 2018 konnte Server Dagstuhl jedoch zu Server Trier aufschließen. Mitlerweile ist Server Dagstuhl der mit Abstand am meisten genutze Server bezüglich der Anzahl der Besucher sowie der Platzierung bei Google.

Insgesamt konnten die Nutzungszahlen in 2022 gegenüber dem Vorjahr leicht gesteigert werden. Insbesondere die Zahl der Seitenaufrufe pro Tag erhöhte sich gegenüber dem Vorjahr um weitere 13,4 Prozent auf nunmehr über 1,49 Millionen Aufrufe pro Tag. Die dblp-Webservices wurden dabei insgesamt von mehr als 850 000 Nutzenden pro Monat besucht. Das entspricht einer Steigerung um 13,8 Prozent.

Fig. 3.4 fasst die durchschnittliche Nutzung aller dblp-Server zusammen. Bei der Bestimmung der unterschiedlichen Nutzenden pro Monat wurde 2021 die Berechnungsgrundlage geändert um zukünftig eine Doppel-Zählung von IP-Adressen auf unterscheidlichen Servern zu vermeiden. Die Zahlen vor 2021 sind daher nur eingeschränkt mit den Folgejahren vergleichbar. Alle Statistiken ignorieren üblicherweise die Zugriffe, die durch bekannte Bot- und Crawler-Software verursacht wurden. Denoch wieß das Nutzungsverhalten in 2022 – insbesondere im Hinblick auf Nutzungs-Sitzungen – eine Reihe von Anomalien auf, die nicht durch das übliche verhalten unserer Nutzenden zu erklären ist. Oftmals traten schon zwischen einzelnen Tagen starken Schwankungen auf. Daher besitzen die Statistiken bezüglich der Nutzungs-Sitzungen in diesem Jahr nur eine recht eingeschränkte Aussagekraft.

In 2022, two official dblp web servers were updated and synchronized on a daily basis:

- server Trier: dblp.uni-trier.de
- server Dagstuhl: dblp.dagstuhl.de

The main domain dblp.org is used as an alias for dblp server Dagstuhl.

In the past, Trier had been the most widely known server due to its high visibility and prominence in the Google search engine. However, during the course of 2018, server Dagstuhl has become increasingly more visible. In 2020, server Dagstuhl overtook Trier in respect to the number of visitors as well as the Google search ranking and has since been the main dblp server.

Overall, the total usage in 2022 slightly increased. In particular, the dblp websites receive on average more than 1.49 million page views per day. That is an increase by 13.4% compared with the previous year. In total, more than 850,000 distinct users from all over the world did use the dblp web services on a monthly basis. This corresponds to an increase by 13.8%.

Figure 3.4 shows the average usage of all three servers. In 2021, the counting method to determine distinct users has been changed in order to avoid double counting of the same IP across diferent servers. Figures from before and after 2021 are therefore only comparable with each other to a limited extent. All figures should ignore the traffic caused by known bots and crawlers. However, user behavior in 2022 – especially with regard to user sessions – showed a number of anomalies that cannot be explained by the usual behavior of our users. Often there were already strong fluctuations between single days. Therefore, the statistics regarding user sessions in this year are not necessarily reliable.

4 Dagstuhl Publishing
Dagstuhl Publishing

Dagstuhl Publishing

Portfolio

4.1

Portfolio

Die Open-Access-Verlagsdienstleistungen von Schloss Dagstuhl werden in der Wissenschaftsgemeinde gut aufgenommen. Im Portfolio des Angebots gibt es zum einen Publikationsserien, die sich auf Veranstaltungen beziehen, die auf Schloss Dagstuhl abgehalten wurden (*Dagstuhl Reports*, *Dagstuhl Manifestos*, *Dagstuhl Follow-Ups*), zum anderen Serien, die Konferenzen und Workshops außerhalb von Schloss Dagstuhl bedienen (*LIPIcs* und *OASIcs*). Ergänzt wird das Portfolio um die wissenschaftliche Zeitschrift *LITES* und die Serie *DARTS*, in der Forschungsartefakte veröffentlicht werden.

Dagstuhl Reports

Alle Dagstuhl-Seminare und Dagstuhl-Perspektiven-Workshops werden in der Zeitschrift *Dagstuhl Reports*[15] dokumentiert, um eine Zitation der Seminare im wissenschaftlichen Kontext zu ermöglichen. Zudem bietet sie auch den Wissenschaftlern, die nicht am Seminar teilgenommen haben, einen zeitnahen Einblick in das, was beim Seminar diskutiert und erarbeitet wurde.

Die Zeitschrift erscheint seit 2011 und enthält in monatlichen Ausgaben Berichte zu den Dagstuhl-Seminaren und -Perspektiven-Workshops, die im jeweiligen Monat stattgefunden haben. Der Inhalt der Berichte wird nicht begutachtet. Das wissenschaftliche Direktorium (siehe Fig. 11.4) agiert als Herausgebergremium für die Reihe. Um umfassende Zusammenstellungen von begutachteten Artikeln auf Basis eines Dagstuhl-Seminars oder -Perspektiven-Workshops zu ermöglichen, wurde die Buchreihe *Dagstuhl Follow-Ups* (siehe unten) gegründet.

In 2022 wurde für 53 Dagstuhl-Seminare und -Perspektiven-Workshops ein Bericht in der Reihe *Dagstuhl Reports* veröffentlicht. An dieser Stelle bedanken wir uns ganz herzlich bei den Organisatoren und Kollektoren für die erfolgreiche Zusammenarbeit.

Dagstuhl Follow-Ups

Die Buchreihe *Dagstuhl Follow-Ups*[16] ermöglicht die Veröffentlichung einer Sammlung begutachteter Beiträge, die auf einem Dagstuhl-Seminar oder Dagstuhl-Perspektiven-Workshop basiert. Für jedes Buch ist ein Antrag notwendig, der vom wissenschaftlichen Direktorium (welches als Herausgebergremium verantwortlich ist) begutachtet und freigegeben werden muss. In 2022 wurde kein Buch in der Reihe veröffentlicht.

[15] https://www.dagstuhl.de/dagrep
[16] https://www.dagstuhl.de/dfu

The scientific community appreciates the Open Access publishing services offered by Schloss Dagstuhl. The portfolio covers series related to events at Schloss Dagstuhl (*Dagstuhl Reports*, *Dagstuhl Manifestos*, *Dagstuhl Follow-Ups*) and series for conferences and workshops held outside of Schloss Dagstuhl (*OASIcs* and *LIPIcs*). The portfolio is supplemented by the scholarly journal *LITES* and by the DARTS series which aims at publishing research artifacts.

Dagstuhl Reports

All Dagstuhl Seminars and Dagstuhl Perspectives Workshops are documented in the periodical *Dagstuhl Reports*[15] which enables the citation of the seminars in a scientific context. Furthermore, it allows scientists who were not able to attend the seminar to inform themselves about the work and discussions of the seminar in a timely manner.

The periodical started with the first seminars of January 2011 and publishes, in monthly issues, reports on Dagstuhl Seminars and Perspectives Workshops that took place in a given month. The content is not peer-reviewed. The Scientific Directorate (see Fig. 11.4) acts as editorial board. For comprehensive collections of peer-reviewed articles developed on the basis of a Dagstuhl Seminar or Perspectives Workshop, we offer seminar organizers the possibility of publishing a volume in our book series *Dagstuhl Follow-Ups* (see below).

In 2022, 53 reports of Dagstuhl Seminars and Dagstuhl Perspectives Workshops have been published. We would like to take this opportunity to cordially thank all organizers and collectors for their successful collaboration.

Dagstuhl Follow-Ups

The *Dagstuhl Follow-Ups*[16] book series is devoted to peer-reviewed collections of original research works that are rooted in a dedicated Dagstuhl Seminar or Dagstuhl Perspectives Workshop. Each book requires a proposal, which is reviewed and finally approved by the Scientific Directorate (which is in charge as editorial board). In 2022, no volume was published in the series.

Dagstuhl Manifestos

Seit 2011 werden in der Zeitschrift *Dagstuhl Manifestos*[17] die Manifestos der Dagstuhl-Perspektiven-Workshops – deren Erstellung zur Aufgabe des Dagstuhl-Perspektiven-Workshops gehört – Open Access veröffentlicht. Das wissenschaftliche Direktorium (siehe Fig. 11.4) fungiert hier ebenfalls als Herausgebergremium. In 2022 wurde kein Buch in der Reihe veröffentlicht

DARTS: Dagstuhl Artifacts Series

In der Reihe *DARTS*[18] werden qualitätsgesicherte Forschungsdaten und -artefakte veröffentlicht. Die Reihe hat dabei die Struktur einer Zeitschrift. In 2022 wurde die siebte Ausgabe mit zwei Heften und insgesamt 33 Artefakten veröffentlicht.

Die Veröffentlichung und Bereitstellung von Forschungdaten und -artefakten ist aktuell ein wichtiges Thema in den wissenschaftlichen Disziplinen und bei den Forschungsfördereinrichtungen. Im Bereich der Informatik wird dieses Thema ebenfalls diskutiert. Im Jahr 2015 gab es zum Beispiel einen Perspektiven-Workshop mit dem Titel „Artifact Evaluation for Publications“[19], der 2016 durch zwei Seminare ergänzt wurde: „Reproducibility of Data-Oriented Experiments in e-Science“[20] und „Rethinking Experimental Methods in Computing“[21].

Schloss Dagstuhl unterstützt mit DARTS die Wissenschaftsgemeinde in der Informatik bei dem Wunsch, Forschungsdaten und -artefakte in einer geeigneten Reihe zu veröffentlichen. Hierbei berücksichtigt DARTS insbesondere auch die Publikationskultur in der Informatik mit ihrem Schwerpunkt auf Konferenzbandveröffentlichungen.

OASIcs: OpenAccess Series in Informatics

Die *OASIcs*-Reihe[22] veröffentlicht begutachtete Tagungsbände von Workshops, Symposien und Konferenzen. Das Herausgebergremium (Fig. 4.1), diskutiert sorgfältig alle Anträge, um ausschließlich qualitativ hochwertige sowie professionell durchgeführte Veranstaltungen in die Reihe aufzunehmen und um gegebenenfalls Empfehlungen zur Verbesserung der Veranstaltungsstruktur zu geben.

In 2022 wurden 10 Bände von thematisch breit gestreuten Workshops und Konferenzen veröffentlicht, siehe Fig. 4.6.

Dagstuhl Manifestos

Since 2011 we have published the manifestos – an expected result of Dagstuhl Perspectives Workshops – in the journal *Dagstuhl Manifestos*[17] in an Open Access manner. The Scientific Directorate (see Fig. 11.4) acts as the editorial board of the journal. In 2022, no volume was published in the series.

DARTS: Dagstuhl Artifacts Series

The *DARTS* series[18] publishes evaluated research data and artifacts. It is organized as a periodical. In 2022, the seventh volume containing two issues with 33 artifacts in total was published.

The publishing of research data and artifacts is currently in the general focus of the scientific community and funding agencies. In the area of computer science, this topic is also under discussion. For example, in 2015 a Perspectives Workshop on “Artifact Evaluation for Publications”[19] took place which was complemented with two seminars in 2016: “Reproducibility of Data-Oriented Experiments in e-Science”[20] and “Rethinking Experimental Methods in Computing”[21].

With DARTS, Schloss Dagstuhl is aiming to support the computing research community with a publishing venue dedicated to research data and artifacts. Especially, DARTS takes into account the publication culture in computer science which focuses on conference proceedings publications.

OASIcs: OpenAccess Series in Informatics

The *OASIcs* series[22] aims to publish the peer-reviewed proceedings of workshops, symposia, and conferences. The editorial board, see Fig. 4.1, discusses carefully all submitted proposals to ensure that only significant and professionally organized events are added to the series and that – if applicable – suggestions are given for improving the structure of the event.

In 2022, Dagstuhl published 10 *OASIcs* volumes covering the proceedings of topically widespread workshops and conferences; see Fig. 4.6.

[17] https://www.dagstuhl.de/dagman
[18] https://www.dagstuhl.de/darts
[19] https://www.dagstuhl.de/15452
[20] https://www.dagstuhl.de/16041
[21] https://www.dagstuhl.de/16111
[22] https://www.dagstuhl.de/oasics

Fig. 4.1
OASIcs Editorial Board.

LIPIcs: Leibniz International Proceedings in Informatics

Die *LIPIcs*-Reihe[23] veröffentlicht Tagungsbände von international renommierten Informatik-Konferenzen, die in ihrem jeweiligen Gebiet führend sind. Das internationale Herausgebergremium (siehe Fig. 4.2) besteht aus einschlägig bekannten Wissenschaftlern und wird seit Oktober 2017 von Luca Aceto als Hauptherausgeber geleitet.

Die Amtszeiten von Christel Baier und Javier Esparza sind 2022 ausgelaufen. Beie haben als langjährige Mitglieder des Herausgebergremiums eine wichtige Rolle in der Entwicklung der Serie gespielt. Für diese Verdienste möchten wir uns an dieser Stelle herzlich bedanken. Christel Baier und Javier Esparza wurden in einem anonymen Wahlverfahren innerhalb des Herausgebergremiums für weitere 4 Jahre in das Gremium gewählt. Siehe auch Fig. 4.2.

In 2022 wurden 36 Tagungsbände von Konferenzen mit insgesamt 1444 Artikeln veröffentlicht; siehe Fig. 4.7 und 4.8.

Auch im zurückliegenden Jahr 2022 gab es wieder viele Anträge bei LIPIcs, womit die große Nachfrage aus den Vorjahren fortgesetzt wurde. In Fig. 4.3 sind alle Konferenzen aufgelistet, deren Anträge 2022 bei LIPIcs positiv begutachtet wurden und mit denen daher eine mehrjährige Kooperation (typischerweise 5 Jahre) eingegangen wurde. Eine dieser Konferenz hat erstmals einen Antrag bei LIPIcs gestellt. Die anderen Konferenzen haben bereits vorher mit LIPIcs kooperiert.

LIPIcs: Leibniz International Proceedings in Informatics

The *LIPIcs* series[23] publishes proceedings of leading conferences in the area of informatics. An international editorial board of renowned researchers (see Fig. 4.2) supervises the conferences that are accepted for LIPIcs and is headed since October 2017 by Luca Aceto.

The terms of Christel Baier und Javier Esparza ended in 2022. Both served as members of the editorial board for several years and played an important role for the development of the series. We would like to take this opportunity to thank them for their extraordinary dedication. Christel Baier und Javier Esparza were voted in anonymous voting within the editorial board for another 4 years membership. See also Fig. 4.2.

The series published 36 proceedings of major conferences with 1444 articles in total in 2022; see Fig. 4.7 and 4.8.

Harvesting the fruits of our long-lasting efforts to attract major conferences to LIPIcs, the year 2022 has again seen several applications for LIPIcs, continuing the high interest from the previous years. Fig. 4.3 lists all conferences that have been accepted in 2022 for a cooperation covering several years (typically 5 years). One of these conferences has submitted a proposal to LIPIcs for the first time. The other conferences have already cooperated with LIPIcs in the past.

[23] https://www.dagstuhl.de/lipics

Fig. 4.2
LIPIcs Editorial Board.

CALCO | Conference on Algebra and Coalgebra in Computer Science accepted for 2023-2025 (Re-Evaluation)
DISC | International Symposium on Distributed Computing accepted for 2022-2026 (Re-Evaluation)
FORC | Symposium on Foundations of Responsible Computing accepted for 2023-2027 (Re-Evaluation)
GD | Graph Drawing and Network Visualization accepted for 2024-2028
GIScience | International Conference on Geographic Information Science accepted for 2023-2027 (Re-Evaluation)
ITC | Conference on Information-Theoretic Cryptography accepted for 2023-2027 (Re-Evaluation)
SEA | International Symposium on Experimental Algorithms accepted for 2023-2027 (Re-Evaluation)
SWAT | Scandinavian Symposium and Workshops on Algorithm Theory accepted for 2022-2026 (Re-Evaluation)
TIME | International Symposium on Temporal Representation and Reasoning accepted for 2022-2024 (Re-Evaluation)
TQC | Conference on the Theory of Quantum Computation, Communication and Cryptography accepted for 2023-2027 (Re-Evaluation)
WABI | International Workshop on Algorithms in Bioinformatics accepted for 2022-2024 (Re-Evaluation)

Fig. 4.3
Conferences accepted in 2022 for publication in LIPIcs.

LITES: Leibniz Transactions on Embedded Systems

Die Open Access-Fachzeitschrift *LITES*[24] veröffentlicht begutachtete Beiträge zu allen Aspekten eingebetteter Systeme. Die Zeitschrift wurde 2012 gegründet und nahm 2013 ihren Betrieb auf. Ein breit aufgestelltes Team an erfahrenen Wissenschaftlern, die für ihr jeweiliges Fachgebiet verantwortlich zeichnen (siehe Fig. 4.4), begutachtet alle eingereichten Arbeiten. Die Zeitschrift wird gemeinsam mit der Fachgruppe *EMbedded Systems Special Interest Group (EMSIG)*[25] der Fachgesellschaft *European Design and Automation Association (EDAA)*[26] herausgege-

LITES: Leibniz Transactions on Embedded Systems

The *LITES*[24] journal publishes original peer-reviewed articles on all aspects of embedded computer systems via Open Access. The journal was established in 2012 and started operating in early 2013. A broad team of experienced researchers, acting as editorial board (see Fig. 4.4), reviews all submitted contributions. The journal is jointly published with the *EMbedded Systems Special Interest Group (EMSIG)*[25] of the *European Design and Automation Association (EDAA)*[26]. The special interest group is responsible for appointing the editorial board,

Dagstuhl Publishing

ben. Die Fachgruppe ist dabei für die Besetzung des Herausgebergremiums verantwortlich, während Schloss Dagstuhl die administrativen Aufgaben der Herausgeberschaft übernimmt.

2022 ist die Amtszeit von Alan Burns zu Ende gegangen. Alan Burns hat als Hauptredakteur seit der Gründung des Journals eine wichtige Rolle bei der Entwicklung der Zeitschrift gespielt.Wir möchten ihm an dieser Stelle für sein außerordentliches Engagement danken.

Björn B. Brandenburg wurde in enger Abstimmung zwischen EDAA/EMSIG und Schloss Dagstuhl als Nachfolger von Alan Burns als neuer Chefredakteur für 5 Jahre gewählt.

In 2022 wurde zwei Ausgaben von *LITES* mit insgesamt 10 Artikeln veröffentlicht.

while Schloss Dagstuhl takes over the administrative tasks of the publication.

In 2022, the term of office of Alan Burns came to an end. Alan Burns has played an important role in the development of the journal as the Editor-in-Chief since its launch.We would like to take this opportunity to thank him for his extraordinary commitment.

Björn B. Brandenburg was appointed in close consultation between EDAA/EMSIG and Schloss Dagstuhl to succeed Alan Burns as the new Editor-in-Chief for a 5-year term.

In 2022, two issue of *LITES* containing 10 articles in total were published.

[24] https://www.dagstuhl.de/lites
[25] http://www.emsig.net/
[26] https://www.edaa.com/

Prof. Alan Burns, DPhil | tenure ended in 2022 University of York, UK | Editor-in-Chief until 2022	Prof. Dr. Martin Fränzle Universität Oldenburg, DE
Dr. Björn B. Brandenburg | tenure started in 2022 MPI-SWS - Kaiserslautern, DE | Editor-in-Chief since 2022	Prof. Steve Goddard University of Iowa - Iowa City, US
Prof. Bashir M. Al-Hashimi University of Southampton, GB	Prof. Dr. Gernot Heiser UNSW - Sydney, AU
Prof. Dr. Karl-Erik Arzén Lund University, SE	Prof. Dr. Axel Jantsch TU Wien, AT
Prof. Neil Audsley City - University of London, GB	Prof. Dr. Sang Lyul Min Seoul National University, KR
Prof. Sanjoy Baruah Washington University, US	Prof. Dr.-Ing. Lothar Thiele ETH Zürich, CH
Prof. Dr. Samarjit Chakraborty University of North Carolina at Chapel Hill, US	Dr. Virginie Wiels ONERA - Toulouse, FR
Prof. Dr.-Ing. Marco Di Natale Sant'Anna School of Advanced Studies - Pisa, IT	

Fig. 4.4
LITES Editorial Board.

Dagstuhl Publishing

Infrastruktur 4.2 *Infrastructure*

Indizierung

Alle Reihen des Publikations-Portfolios werden bei *dblp* gelistet, siehe Fig. 4.5. Die Bände aus den Reihen *LIPIcs* und *OASIcs* werden zudem bei Scopus[27] eingereicht, wo sie regelmäßig indiziert werden. Die Reihen *LIPIcs* und *OASIcs* sowie die Zeitschrift *LITES* sind zudem im Directory of Open Access Journals (DOAJ) gelistet, siehe Fig. 4.5. Zudem unterstützen die technischen Schnittstellen die Datenakquise durch GoogleScholar, was eine gute Sichtbarkeit der Publikationen garantiert.

LeibnizOpen

Die Leibniz-Gemeinschaft hat mit *LeibnizOpen*[28] ein Online-Repositorium ins Leben gerufen, um Open Access-Veröffentlichungen von Leibniz-Instituten und deren Wissenschaftlern zu unterstützen und sichtbar zu machen. Schloss Dagstuhl liefert alle Artikel aus den Reihen *Dagstuhl Reports* und *Dagstuhl Manifestos* an das Repositorium und stärkt dadurch Forschungsergebnisse aus der Informatik innerhalb dieses multidisziplininären Repositoriums.

Indexing

All series of the publication portfolio are listed in *dblp*; see Fig. 4.5. The LIPIcs and OASIcs volumes are submitted to Scopus[27] where they are regularly indexed. The *LIPIcs* and *OASIcs* series as well as the journal *LITES* are also listed in the Directory of Open Access Journals (DOAJ), see Fig. 4.5. The technical interface of our publication server enables harvesting according to the Google Scholar guidelines. Google Scholar regularly retrieves metadata and full-texts from our server.

LeibnizOpen

The Leibniz Association has established the *LeibnizOpen*[28] repository to promote the open-access publications of Leibniz institutes and their researchers. Schloss Dagstuhl submits all articles from the *Dagstuhl Reports* and *Dagstuhl Manifestos* series to the repository, thereby strengthening informatics-related research in this multi-disciplinary repository.

[27] https://www.scopus.com
[28] http://www.leibnizopen.de/

dblp
Dagstuhl Reports https://dblp.org/db/journals/dagstuhl-reports/
Dagstuhl Manifestos https://dblp.org/db/journals/dagstuhl-manifestos/
Dagstuhl Follow-Ups https://dblp.org/db/series/dfu/
OASIcs https://dblp.org/db/series/oasics/
LIPIcs https://dblp.org/db/series/lipics/
LITES https://dblp.org/db/journals/lites/
DARTS https://dblp.org/db/journals/darts/

DOAJ
OASIcs https://doaj.org/toc/2190-6807
LIPIcs https://doaj.org/toc/1868-8969
LITES https://doaj.org/toc/2199-2002

Fig. 4.5
Indexing of Dagstuhl Publishing series in dblp and DOAJ.

Dagstuhl Publishing

Engagement in Arbeitsgruppen

Mitarbeitende von Schloss Dagstuhl waren auch in 2022 in verschiedenen Arbeitsgruppen beteiligt:

- Projektgruppe "Open Science-Leitbild" der Leibniz-Gemeinschaft
- Arbeitsgruppe "Open Access und Publikationsmanagement" der Leibniz-Gemeinschaft
- Arbeitsgruppe "Open Access" des Saarlandes
- Task Force „Research Software Infrastructures" der EOSC (European Open Science Cloud)
- Arbeitsgruppe "Forschungsdaten" der Leibniz-Gemeinschaft

Publikationsserver: DROPS

Über den Dagstuhl Research Online Publication Server (DROPS)[29] werden alle Veröffentlichungen von Schloss Dagstuhl verwaltet. Es werden hierbei die allgemeinen Richtlinien für Online-Publikationen gemäß der Dublin Core-Initiative[30] berücksichtigt, insbesondere werden alle nötigen Metadaten zu jeder Publikation gespeichert und die Langzeitverfügbarkeit sichergestellt.

Einreichungssystem: DSub

Im Frühjahr 2019 wurde ein von Dagstuhl entwickeltes Einreichungssystem namens DSub in Betrieb genommen. Mit diesem System werden seitdem alle Einreichungen zu den Reihen LIPIcs und OASIcs entgegengenommen. Unter anderem wurde mit diesem System dem Wunsch einer aktiven Autorenfreigabe der überarbeiteten Dokumente vor der Veröffentlichung entsprochen und die automatische Extraktion der Metadaten aus den LaTeX-Quellen ermöglicht.

Langzeitarchivierung

Alle Publikationen werden bei der Deutschen Nationalbibliothek (DNB)[31] zur (digitalen) Langzeitarchivierung eingereicht.

Mirroring

Um dem Verlust von Daten vorzubeugen, werden seit 2010 zwei Kooperationen zur Spiegelung (Mirroring) von Inhalten des Publiktionsservers DROPS gepflegt:

- emis.de: Das unter Leitung des FIZ Karlsruhe, Leibniz-Institut für Informationsinfrastruktur, organisierte Mathematik-Publikations-Portal European Mathematical Information Service (EMIS) spiegelt alle Bände der LIPIcs-Reihe.[32]
- SunSite Central Europe: Der gesamte DROPS-Bestand wird in regelmäßigen Abständen auf der SunSite Aachen gespiegelt.[33]

Involvement in working groups

Employees of Schloss Dagstuhl were also involved in 2022 in various working groups:

- Project group "Open Science-Leitbild (engl. Open Science-mission statement)" of the Leibniz association
- Working group "Open Access and Publication Management" of the Leibniz association
- Working group "Open Access" of Saarland
- Research Software Infrastructures Task Force of the EOSC (European Open Science Cloud)
- Working group "Research Data" of the Leibniz association

Publication Server: DROPS

All items published by the center are administered via the Dagstuhl Research Online Publication Server (DROPS)[29]. The general guidelines of the Dublin Core initiative[30] applicable to online publications are adhered to, meaning that all the requisite metadata of each publication is stored, thus ensuring availability in the long term.

Submission System: DSub

In spring 2019 a submission system called DSub developed by Dagstuhl was introduced. Since then, this system has been used to process all submissions for the LIPIcs and OASIcs series. Among other things, the new system has satisfied the need for active author approval of revised documents prior the publication and enables automatic extraction of metadata from LaTeX sources.

Long-Term Archiving

All publications are submitted to the German National Library (DNB)[31] for (digital) long-term archiving.

Mirroring

In order to prevent data loss, two cooperative ventures were initiated in 2010 for mirroring the content of the DROPS publication server:

- emis.de: The portal for electronic math resources European Mathematical Information Service (EMIS), organized under the auspices of FIZ Karlsruhe – Leibniz Institute for Information Infrastructure, mirrors all volumes of the LIPIcs series[32].
- SunSite Central Europe: All the DROPS assets are mirrored at regular intervals on the Aachen SunSite.[33]

[29] https://www.dagstuhl.de/drops
[30] http://dublincore.org/
[31] https://www.dnb.de/DE/Professionell/Erhalten/erhalten_node.html#sprg209698
[32] https://subs.emis.de/LIPIcs/
[33] http://vesta.informatik.rwth-aachen.de/Dagstuhl/

Vol. 97 | 3rd International Conference on Blockchain Economics, Security and Protocols (Tokenomics 2021)
https://www.dagstuhl.de/dagpub/978-3-95977-220-4

Vol. 98 | Third Workshop on Next Generation Real-Time Embedded Systems (NG-RES 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-221-1

Vol. 99 | International Research School in Artificial Intelligence in Bergen (AIB 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-228-0

Vol. 100 | 13th Workshop on Parallel Programming and Run-Time Management Techniques for Many-Core Architectures and 11th Workshop on Design Tools and Architectures for Multicore Embedded Computing Platforms (PARMA-DITAM 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-231-0

Vol. 101 | 5th International Symposium on Foundations and Applications of Blockchain 2022 (FAB 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-248-8

Vol. 102 | Third International Computer Programming Education Conference (ICPEC 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-229-7

Vol. 103 | 20th International Workshop on Worst-Case Execution Time Analysis (WCET 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-244-0

Vol. 104 | 11th Symposium on Languages, Applications and Technologies (SLATE 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-245-7

Vol. 105 | 4th International Workshop on Formal Methods for Blockchains (FMBC 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-250-1

Vol. 106 | 22nd Symposium on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-259-4

Fig. 4.6
OASIcs volumes published in 2022.

Vol. 215 | 13th Innovations in Theoretical Computer Science Conference (ITCS 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-217-4

Vol. 216 | 30th EACSL Annual Conference on Computer Science Logic (CSL 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-218-1

Vol. 217 | 25th International Conference on Principles of Distributed Systems (OPODIS 2021)
https://www.dagstuhl.de/dagpub/978-3-95977-219-8

Vol. 218 | 3rd Symposium on Foundations of Responsible Computing (FORC 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-226-6

Vol. 219 | 39th International Symposium on Theoretical Aspects of Computer Science (STACS 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-222-8

Vol. 220 | 25th International Conference on Database Theory (ICDT 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-223-5

Vol. 221 | 1st Symposium on Algorithmic Foundations of Dynamic Networks (SAND 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-224-2

Vol. 222 | 36th European Conference on Object-Oriented Programming (ECOOP 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-225-9

Vol. 223 | 33rd Annual Symposium on Combinatorial Pattern Matching (CPM 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-234-1

Vol. 224 | 38th International Symposium on Computational Geometry (SoCG 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-227-3

Vol. 225 | 33rd International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-230-3

Vol. 226 | 11th International Conference on Fun with Algorithms (FUN 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-232-7

Fig. 4.7
LIPIcs volumes published in 2022 – Part 1.

Vol. 227 | 18th Scandinavian Symposium and Workshops on Algorithm Theory (SWAT 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-236-5

Vol. 228 | 7th International Conference on Formal Structures for Computation and Deduction (FSCD 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-233-4

Vol. 229 | 49th International Colloquium on Automata, Languages, and Programming (ICALP 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-235-8

Vol. 230 | 3rd Conference on Information-Theoretic Cryptography (ITC 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-238-9

Vol. 231 | 34th Euromicro Conference on Real-Time Systems (ECRTS 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-239-6

Vol. 232 | 17th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-237-2

Vol. 233 | 20th International Symposium on Experimental Algorithms (SEA 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-251-8

Vol. 234 | 37th Computational Complexity Conference (CCC 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-241-9

Vol. 235 | 28th International Conference on Principles and Practice of Constraint Programming (CP 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-240-2

Vol. 236 | 25th International Conference on Theory and Applications of Satisfiability Testing (SAT 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-242-6

Vol. 237 | 13th International Conference on Interactive Theorem Proving (ITP 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-252-5

Vol. 238 | 28th International Conference on DNA Computing and Molecular Programming (DNA 28)
https://www.dagstuhl.de/dagpub/978-3-95977-253-2

Vol. 239 | 27th International Conference on Types for Proofs and Programs (TYPES 2021)
https://www.dagstuhl.de/dagpub/978-3-95977-254-9

Vol. 240 | 15th International Conference on Spatial Information Theory (COSIT 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-257-0

Vol. 241 | 47th International Symposium on Mathematical Foundations of Computer Science (MFCS 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-256-3

Vol. 242 | 22nd International Workshop on Algorithms in Bioinformatics (WABI 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-243-3

Vol. 243 | 33rd International Conference on Concurrency Theory (CONCUR 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-246-4

Vol. 244 | 30th Annual European Symposium on Algorithms (ESA 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-247-1

Vol. 245 | Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-249-5

Vol. 246 | 36th International Symposium on Distributed Computing (DISC 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-255-6

Vol. 247 | 29th International Symposium on Temporal Representation and Reasoning (TIME 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-262-4

Vol. 248 | 33rd International Symposium on Algorithms and Computation (ISAAC 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-258-7

Vol. 249 | 17th International Symposium on Parameterized and Exact Computation (IPEC 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-260-0

Vol. 250 | 42nd IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2022)
https://www.dagstuhl.de/dagpub/978-3-95977-261-7

Fig. 4.8
LIPIcs volumes published in 2022 – Part 2.

5 Resonanz
Feedback

Resonanz zu Seminaren und Workshops

5.1

Feedback on Seminars and Workshops

Resonanz von Teilnehmern

Schloss Dagstuhl bekommt viel Lob von seinen Gästen, meistens in mündlicher Form, wenn die Gäste nach einer intensiven Seminarwoche das Schloss verlassen. Manche Gäste nehmen sich jedoch auch die Zeit, uns nachträglich zu schreiben und ihre Eindrücke mit uns zu teilen. Im Folgenden haben wir mit freundlicher Genehmigung der Autoren einen Auszug aus unserer großen Sammlung an Dankeschön-Nachrichten zusammengestellt.

Feedback from Participants

Schloss Dagstuhl receives a lot of positive feedback, typically verbally when our guests are checking out after an intense seminar. However, many guests take the time to write to us about their impressions. What follows is an excerpt from our large thank-you collection, cited here with the authors' appreciated permission.

Maria Eugenia Monge

22181 – Computational Metabolomics: From Spectra to Knowledge | Dagstuhl Seminar | https://www.dagstuhl.de/22181

This is just a quick message to let you know that the Dagstuhl seminar I attended last week was a highly constructive and fruitful academic experience for me. One of the best scientific events I have ever attended. Many thanks for all the assistance you provided and for the support I received from Schloss Dagstuhl to make my participation possible.

Resonanz unserer Organisatoren

Der Erfolg von Schloss Dagstuhl hängt im wesentlichen Maße auch von den Seminarorganisatoren ab, die interessante und neue Themen vorschlagen. Wir sind hoch erfreut, dass die Seminarorganisatoren selber die Angebote und die Umgebung, die wir zur Verfügung stellen, schätzen. Im Folgenden geben wir mit freundlicher Genehmigung der Autoren einige der Kommentare unsere Seminarorganisatoren wieder.

Feedback from Organizers

The success of Schloss Dagstuhl depends to a large extent on our outstanding seminar organizers, who continually enrich the scientific program with a range of interesting and new topics. We are very glad to be able to provide services and an environment that organizers appreciate. The following comments from organizers are excerpted from the Dagstuhl Report or personal emails to us. We cite them with their kindly permission.

Organizers of Dagstuhl Seminar 22051

22051 – Finite and Algorithmic Model Theory | Dagstuhl Seminar | https://www.dagstuhl.de/22051

The organizers are grateful to the Scientific Directorate of the Center for its support of this workshop and the staff of Schloss Dagstuhl for their organisation of our stay (including regular covid testing) and their hospitality, despite the many challenges posed by covid.

Organizers of Dagstuhl Seminar 22021

22021 – Mobility Data Science | Dagstuhl Seminar | https://www.dagstuhl.de/22021

In the time of COVID-19, organizing a scientific gathering is challenged by the uncertainty of physical participation. The organizer would like to thank the Dagstuhl team, which allowed the switching from purely onsite into hybrid only a few weeks before the seminar due to the Omicron variant of COVID-19. This flexibility, which is not matched by hotels and conference centers, allowed the organizers to focus on the program and the participation, rather than focusing on the venue logistics.

Organizers of Dagstuhl Seminar 22082

22082 – Deep Learning and Knowledge Integration for Music Audio Analysis | Dagstuhl Seminar | https://www.dagstuhl.de/22082

Many of our participants were visiting Dagstuhl for the first time and enthusiastically praised the open and inspiring setting. The group dynamics were excellent, with many personal exchanges and shared activities. Some scientists expressed their appreciation for having the opportunity for prolonged discussions with researchers from neighboring research fields, which is often impossible during conference-like events. At this point, we would like to let some of the participants have their say:
Stefan Balke (pmOne – Paderborn, DE): "Dagstuhl is always a wonderful experience, having time to think, talk, and play music. All in a relaxed atmosphere, the seminar feels like a family meeting – especially in these times."
Alice Cohen-Hadria (IRCAM – Paris, FR): "Now I feel like a part of a community."
Dasaem Jeong (Sogang University – Seoul, KR): "Full of insightful discussions, music, and friends in a beautiful place."
Cynthia Liem (TU Delft, NL): "Dagstuhl is the one place in the world where one effectively can have a week long unconference. More deeply talking about research and new ideas, enjoying time with academic friends, with much less distraction than one would have at home, or even in a 'regular' conference. Especially coming out of a pandemic, I am realizing this is among the most valuable things in our professional life."
Daniel Stoller (Spotify – Bonn, DE): "Dagstuhl brings perspectives on the big issues."
Yu Wang (New York University – Brooklyn, US): "Discussion is like music: the live version is always better."
In conclusion, our expectations for the seminar were not only met but exceeded, in particular concerning networking and community building. We want to express our gratitude to the Dagstuhl board for giving us the opportunity to organize this seminar, the Dagstuhl office for their exceptional support in the organization process, and the entire Dagstuhl staff for their excellent service during the seminar.

Resonanz in Sozialen Netzwerken

Mehr und mehr Gäste nutzen die Möglichkeiten des Webs wie Twitter und Blogs um über ihre Erfahrungen in Dagstuhl zu berichten. Wir geben hier einige Referenzen.

Feedback in Social Media

More and more of our guests are using social media such as Twitter and blogs to share their experiences of Dagstuhl with others. Below are some selected excerpts.

Shriram Krishnamurthi (Brown University – Providence, US)

Twitter | https://twitter.com/ShriramKMurthi/status/1488232944711999493

Thought I'd feel more melancholy about my first virtual Dagstuhl; surprised I don't. Also never thought a hybrid event could work well, but this is splendid (and, to me, unlike any virtual *conference* — maybe because we have a focused topic and small group?). Two nice surprises.

Andreas Stefik (University of Nevada – Las Vegas, US)

Twitter | https://twitter.com/AndreasStefik/status/1489045174151892996

social connection is hard. One thing I love about Dagstuhl, just absolutely love, is meeting people I do not know in research and "hanging out in a castle". That's hard to replicate, even if you stayed on zoom for 8 hours. It's just not the same.

Michael Bonfert (Universität Bremen, DE)

Google Maps | https://goo.gl/maps/kzHYcnhpC8x5RkdC8

A wonderful place of encounter and exchange. Beyond what's possible in a quick conference coffee break, you find the time and an ideal environment to connect with fellow researchers to advance the field and come up with new collaboration ideas. The team takes wonderful care of everyone. Also individual requests concerning dietary restrictions or another pillow are considered. Many thanks to everyone involved backstage

Rafael Oliveira (University of Waterloo, CA)
Google Maps | https://goo.gl/maps/7L1EXYxDkcdLXdWv6

Excellent research center! The workshops are wonderful and the center is setup to encourage stimulating research collaboration! I can't wait to visit again!

Marco Lübbecke (RWTH Aachen, DE)
Google Maps | https://goo.gl/maps/5wuEw7TFxhH8vMeT9

the Oberwolfach of computer science, wonderful facilities, hard to reach (reserve a taxi way ahead of time!)

Holger Schlingloff (HU Berlin, DE)
Google Maps | https://goo.gl/maps/6fSQm1iQqRBujbRs7

for computing scientists this is the best place in the world

Resonanz im Fragebogen

Jeder Teilnehmer erhält von uns einen Fragebogen zur Evaluation des vom Teilnehmer besuchten Dagstuhl-Seminars oder Dagstuhl-Perspektiven-Workshops. Durch diese anonyme Befragung erhalten wir ebenfalls eine Menge positiver Kommentare. Im Folgenden zitieren wir hier einige von diesen.

Survey Feedback

Every participant has the opportunity to fill out a questionnaire about the Dagstuhl Seminar or Dagstuhl Perspectives Workshop they attended for evaluation purposes. Below are some excerpts from the many positive comments we received through this anonymous survey.

22021 – Mobility Data Science | Dagstuhl Seminar | https://www.dagstuhl.de/22021

Right now, hybrid seminars are a necessity and clearly superior to not having seminars. But the in-person conversations are so much more useful and rewarding and having to return to your computer and speak to a screen after an inspiring meal is a real drag. So as soon as the pandemic situation allows, Dagstuhl should go back to in-person only. In person discussions are priceless; I'd rather go to one in-person event a year than 10 online or hybrid ones. I have observed, though, that the seminars have ever increasing numbers of participants. That's not a feature (also see above). Smaller groups allow you to get to know everybody (at least to some degree), the huge seminars loose the all important personal connections.

22021 – Mobility Data Science | Dagstuhl Seminar | https://www.dagstuhl.de/22021

I am VERY glad to see Dagstuhl enduring the current daring conditions and being able to provide an environment where researchers can, despite all adversities, come "together" and advance their research agendas. Thank you Dagstuhl!

22052 – The Human Factors Impact of Programming Error Messages | Dagstuhl Seminar | https://www.dagstuhl.de/22052

I confess that I didn't know about Schloss Dagstuhl prior to being invited. I dimly sensed that it was a huge opportunity and honor but it was only after I got here that I realized just how much of an honor it was to be invited. I am very grateful to have been included in this seminar. I loved being here. I have an invitation to come back in October and I am already trying to see if I can make it. On suggestions: Is there any way for us to suggest colleagues whom we think could contribute to the seminar? I realize that the list of attendees is vetted. Is there a way for us to suggest names that may be considered?

22052 – The Human Factors Impact of Programming Error Messages | Dagstuhl Seminar | https://www.dagstuhl.de/22052

Please accept that hybrid is here to stay. Even after the pandemic, cost, visas, etc. remain a problem. They always were, and we ignored them. Now they are being pushed to the forefront – people who suffer from these issues are getting used to having access, and will notice when it is taken away. While I deeply love in-person Dagstuhl, I also want to get the broadest group of experts we can and I want to hear from those voices we don't always get to hear – especially when dealing with human-facing issues, it's *IMPORTANT* to hear from people from different countries/continents who are not normally represented.

22052 – The Human Factors Impact of Programming Error Messages | Dagstuhl Seminar | https://www.dagstuhl.de/22052

There's nothing in CS like Dagstuhl. Its relevance is beyond question.

22072 – New Perspectives in Symbolic Computation and Satisfiability Checking | Dagstuhl Seminar | https://www.dagstuhl.de/22072

This was such a welcome brake from covid, you did a great job in building a "bubble" for us to feel safe. Thanks.

22072 – New Perspectives in Symbolic Computation and Satisfiability Checking | Dagstuhl Seminar | https://www.dagstuhl.de/22072

The ability to have direct contacts is absolutely fundamental. I would not encourage remote attendance, once the pandemic is gone.

22072 – New Perspectives in Symbolic Computation and Satisfiability Checking | Dagstuhl Seminar | https://www.dagstuhl.de/22072

Dagstuhl is great. I think you are doing a great job at preserving what should be preserved while making it evolve.

22082 – Deep Learning and Knowledge Integration for Music Audio Analysis | Dagstuhl Seminar | https://www.dagstuhl.de/22082

I think the format is just right and there are many nice things that should be kept (e.g. random seating in dining room, free wednesday afternoon, and open schedule).

22082 – Deep Learning and Knowledge Integration for Music Audio Analysis | Dagstuhl Seminar | https://www.dagstuhl.de/22082

Dagstuhl should not become a virtual experience like some scientific venue are becoming in my opinion. It is at its very core an in person venue.

22102 – Computational Models of Human-Automated Vehicle Interaction | Dagstuhl Seminar | https://www.dagstuhl.de/22102

Dagstuhl is a world-leading institution that embraces academic diversity with a clear focus on computing. It is highly appreciated in my community, which spans engineering psychology, humanities, and applied computer science.

22111 – Database Indexing and Query Processing | Dagstuhl Seminar | https://www.dagstuhl.de/22111

Love Dagstuhl, always looking forward to visiting it!

22121 – 3D Morphable Models and Beyond | Dagstuhl Seminar | https://www.dagstuhl.de/22121

It is a very nice seminar. Dagstuhl is definitely the ideal place for scholars and experts in industry.

22141 – Symmetric Cryptography | Dagstuhl Seminar | https://www.dagstuhl.de/22141

dagstuhl might want to capitalise in its physical nature as other encounters shift to online attendance

22162 – Urban Mobility Analytics | Dagstuhl Seminar | https://www.dagstuhl.de/22162

Keep the friendly staff. Continue all those lovely wooden benches around the Schloss.

22161 – Recent Advancements in Tractable Probabilistic Inference | Dagstuhl Seminar | https://www.dagstuhl.de/22161

As a general remark, the Dagstuhl staff was even more helpful than at previous events, something I would have not thought it possible before. They were really outstanding!

22181 – Computational Metabolomics: From Spectra to Knowledge | Dagstuhl Seminar | https://www.dagstuhl.de/22181

The ability to de-emphasize giving formal talks and providing a „safe place" to share ideas, including non-published ones, is incredibly fruitful.

22181 – Computational Metabolomics: From Spectra to Knowledge | Dagstuhl Seminar | https://www.dagstuhl.de/22181

Avoid going hybrid by default. It is great to have an opportunity to join individual (!) experts for a videocall in some of the rooms.

22191 – Visual Text Analytics | Dagstuhl Seminar | https://www.dagstuhl.de/22191

Very glad to be invited to Dagstuhl, the professional and social experience is outstanding, and definitely looking forward to the opportunities to return here in the future!

22262 – Human-Centered Artificial Intelligence | Dagstuhl Perspectives Workshop | https://www.dagstuhl.de/22262

Be flexible about the nature of 'outputs'. It is hard to write a paper or a book chapter in 4 days. The real value is about connections and idea sparking. The real benefits are personal and collaborative and long term. These are hard to metricise or prove, but they are there and valuable.

22272 – Eat-IT: Interactive Food | Dagstuhl Seminar | https://www.dagstuhl.de/22272

Please keep it. It is important to have a place where people can meet onsite. To stay a whole week working together is an outstanding experience. This makes the difference between Dagstuhl and the rest of the world.

22292 – Computational Approaches for Digitized Historical Newspapers | Dagstuhl Seminar | https://www.dagstuhl.de/22292

Keep up the good work towards a better gender balance of Dagstuhl participants.

22301 – Algorithmic Aspects of Information Theory | Dagstuhl Seminar | https://www.dagstuhl.de/22301

I want to really thank the organizers for arranging the seminar in a hybrid mode. I have (severe, government-imposed) travel restrictions that prevented me from attending the seminar in person. But I ended up having benefited significantly from the talks, which were truly outstanding. This is my first Dagstuhl workshop and will not be my last for sure.

22331 – Visualization and Decision Making Design Under Uncertainty | Dagstuhl Seminar | https://www.dagstuhl.de/22331

This is one of the most unique and inspiring collaborative experiences I have had. I am from a tangential field of geological sciences. I wish that we had something similar in my field.

22341 – Power and Energy-Aware Computing on Heterogeneous Systems (PEACHES) | Dagstuhl Seminar | https://www.dagstuhl.de/22341

Nowadays where everything is performance-measured: try to keep the deciders convinced that, if not directly, it will ultimately pay off to give researchers a bit of some "free playground".

22351 – Interactive Visualization for Fostering Trust in ML | Dagstuhl Seminar | https://www.dagstuhl.de/22351

nice that you have special Dagstuhl[-labeled] wine which can be bought as a gift,

22351 – Interactive Visualization for Fostering Trust in ML | Dagstuhl Seminar | https://www.dagstuhl.de/22351

Dagstuhl was an incredible experience – to gather diverse researchers with most distractions removed to focus on important problems. The speed of innovation in our discussions is inspiring.

22351 – Interactive Visualization for Fostering Trust in ML | Dagstuhl Seminar | https://www.dagstuhl.de/22351

Fixed lunch and dinner times were helpful –> it helps keep discussions focused when you know there is a deadline by which they must be completed. Although I don't think that hybrid seminars work well for a Dagstuhl Seminar, I would welcome it if Dagstuhl could keep the possibilities to, e.g. invite a guest speaker virtually to one of the sessions if that person otherwise could not participate in the seminar (but this would be just an invitation to give an online talk, not a virtual participation)

22362 – Model-Driven Engineering of Digital Twins | Dagstuhl Seminar | https://www.dagstuhl.de/22362

It is always a pleasure to come to Dagstuhl. It provides a unique experience to focus on research problems at hand and discuss with experts in the field in an atmosphere which is not provided by any other event.

22372 – Knowledge Graphs and Their Role in the Knowledge Engineering of the 21st Century | Dagstuhl Seminar | https://www.dagstuhl.de/22372

It is wonderful to be here, a magic concept to stimulate scientific discourse.

22381 – Rational Design of RiboNucleic Acids | Dagstuhl Seminar | https://www.dagstuhl.de/22381

Thanks for this amazing facility, and congratulation on a seamless yet strict application of health measures in this pandemic.

22391 – Cognitive Robotics | Dagstuhl Seminar | https://www.dagstuhl.de/22391

The office staff were great and were quick to respond to queries. The kitchen staff were very friendly and efficient, especially when serving extra food. The housekeeping staff did an excellent job of cleaning up my room. The meals were very tasty and had a nice range of different dishes. The leisure facilities I tried (pool room, breakout rooms) were great. Overall, excellent service.

22392 – Transparent Quantitative Research as a User Interface Problem | Dagstuhl Seminar | https://www.dagstuhl.de/22392

Just wanted to say that I thought this was the best conference/seminar venue I have ever been to. In particular, I thought it was fantastic that you even had special pillows for people to request!

22392 – Transparent Quantitative Research as a User Interface Problem | Dagstuhl Seminar | https://www.dagstuhl.de/22392

Loved the fitness room as a place of socialization and gaming (e.g., foosball)!

22442 – Toward Scientific Evidence Standards in Empirical Computer Science | Dagstuhl Seminar | https://www.dagstuhl.de/22442

This was the best academic event I've ever attended. I am not being hyperbolic. It was all hits, no misses. I enjoyed every session. It was well organized. The talks were interesting and relevant to my work. I met a bunch of likeminded people who I probably never would have found otherwise because we're in different communities.

22461 – Dynamic Graph Algorithms | Dagstuhl Seminar | https://www.dagstuhl.de/22461

I like the bikes, music room, pool, ping pong and games

22462 – Set Visualization and Uncertainty | Dagstuhl Seminar | https://www.dagstuhl.de/22462

I have been attending these seminars for around 10 years, they are the gold standard for encouraging collaboration between academics. If any of my junior collegues ask if they should accept an invitation, I always say they should prioritize Dagstuhl. Don't change, the system works really well.

22481 – Vertex Partitioning in Graphs: From Structure to Algorithms | Dagstuhl Seminar | https://www.dagstuhl.de/22481

I appreciate that Dagstuhl provides a very good support for those with small children to attend (I understand that this could include arranging a nanny) - indeed, the seminar was attended by a couple with small baby and it was great to see arrangements that were made to permit their smooth participation in the seminar. However, I was surprised that the room in the attic was completely converted to a children room. I believe that the room could serve well both as a discussion room and a children room (when needed for this purpose) if a different layout was implemented - a different layout (e.g., moving the castle to the other side and remounting the whiteboard at its original place) would permit both usages and it seems to me to be a better use of the space.

Resonanz zur Bibliographiedatenbank dblp

5.2

Feedback on the dblp computer science bibliography

Die Bibliographiedatenbank dblp wird von zahlreichen internationalen Wissenschaftlern hoch geschätzt und erhält viel Lob. Feedback erhalten wir per Mail, durch Gespräche mit Forschern vor Ort in Dagstuhl, oder durch die sozialen Medien.

The dblp computer science bibliography is internationally well known and appreciated. We receive a lot of feedback via mail, through discussions with researchers at Schloss Dagstuhl, and via social media.

Resonanz in Sozialen Netzwerken

Social Media Feedback

Antonio Vergari, University of Edinburgh, UK

Twitter | https://twitter.com/tetraduzione/status/1519299704156311553

The Max Planck Society (@maxplanckpress) is super cool and all [...], but the Leibniz Association (@LeibnizWGL) is giving us Dagstuhl Seminars (@dagstuhl) and DBLP (@dblp_org)!

John Wickerson, Imperial College London, UK

Twitter | https://twitter.com/wicko3/status/1504128165769273344

A nice touch by the @DateConference organisers: they have emailed me to suggest some papers I might like to attend at the conference next week, based on my DBLP profile.

Vasileios Lampos, University College London, UK

Twitter | https://twitter.com/lampos/status/1483887841600053264

DBLP offers an incomplete coverage to say the least (CS is very much an interdisciplinary endeavour nowadays), Semantic Scholar is a complex mess (how does account/author merging work?), and Google Scholar is BY FAR the best (but far from perfect).

Tim Vieira, Johns Hopkins University, MD, USA

Twitter | https://twitter.com/xtimv/status/1507399642186719260

Pro-tip: DBLP will give you significantly better BibTeX than any other site. Copy-pasting BibTeX from Google Scholar should be outlawed.

Houcemeddine Turki, University of Sfax, Tunisia

Twitter | https://twitter.com/Csisc1994/status/1549861423790030850

Best practices for the Coalition for Reforming Research Assessment should involve the use of open resources such as @opencitations, @dblp_org, @OpenAIRE_eu, @wikidata, @CrossrefOrg, @ORCID_Org, and @orkg_org for a transparent and reproducible research evaluation.

Brian Hempel, University of Chicago, IL, USA

Twitter | https://twitter.com/brianhempel/status/1557209707261075457

I have a command line script that takes any CS paper title and turns it into clean Bibtex via DBLP. I can cite papers super fast. I'm not sure how I would hit deadlines without it.

Clément L. Canonne, University of Sydney, Australia

Twitter | https://twitter.com/ccanonne_/status/1556129669140713472

Uncontroversial take: DBLP is a wonderful and indispensable service, and the CS community as a whole is so lucky it exists.

Denilson Barbosa, University of Alberta, Canada

Twitter | https://twitter.com/ddmbarbosa/status/1579876296615940097

One of my grad students just asked me if I knew about this amazingly clean and useful bibliography resource called @dblp_org. Yup, I knew about it. :)

Sam Tobin-Hochstadt, Indiana University, IN, USA

Twitter | https://twitter.com/samth/status/1582155142467817476

I find that the dblp bibtex is much better than the publisher bibtex.

Steve Blackburn, Google Brain and Australian National University, Australia

Twitter | https://twitter.com/stevemblackburn/status/1582208754220552193

As far as I know the online source that is most comprehensive and sucks least is DBLP. I systematically use that as my starting point.

Dominic Steinhöfel, CISPA Helmholtz Center for Information Security, Germany

Twitter | https://twitter.com/DomSteinhoefel/status/1588188595336142852

Use DBLP! Many problems can be avoided using the correct sources for BibTeX code. In my experience, https://dblp.org is *the* best source for computer science BibTeX material available on the Internet.

Gautam Kamath, University of Waterloo, Canada

Twitter | https://twitter.com/thegautamkamath/status/1597427260721463296

Wanted to give a shoutout to the amazing resource @dblp_org. I emailed them on Sunday, about a bunch of ambiguous entries for various faculty. This evening (Monday), they had quietly fixed all of them. What an amazing service. I take it for granted too often.

6 Die Seminare in 2022
The 2022 Seminars

Applications, Interdisciplinary Work

- Algorithms for Participatory Democracy (22271)
- Anticipatory Human-Machine Interaction (22202)
- Challenges and Opportunities of Democracy in the Digital Society (22361)
- Cognitive Augmentation (22491)
- Computational Approaches for Digitized Historical Newspapers (22292)
- Computational Metabolomics: From Spectra to Knowledge (22181)
- Computational Models of Human-Automated Vehicle Interaction (22102)
- Computer Science Methods for Effective and Sustainable Simulation Studies (22401)
- Dynamic Traffic Models in Transportation Science (22192)
- Eat-IT: Interactive Food (22272)
- Exponential Analysis: Theoretical Progress and Technological Innovation (22221)
- Framing in Communication: From Theories to Computation (22131)
- Human-Game AI Interaction (22251)
- Mobility Data Analysis: From Technical to Ethical (22022)
- Mobility Data Science (22021)
- Radical Innovation and Design for Connected and Automated Vehicles (22222)
- Rational Design of RiboNucleic Acids (22381)
- Technologies to Support Critical Thinking in an Age of Misinformation (22172)
- Toward Scientific Evidence Standards in Empirical Computer Science (22442)
- Transparent Quantitative Research as a User Interface Problem (22392)
- Urban Mobility Analytics (22162)

Artificial Intelligence, Machine Learning, Computational Linguistics

- AI for the Social Good (22091)
- Cognitive Robotics (22391)
- Current and Future Challenges in Knowledge Representation and Reasoning (22282)
- Deep Learning and Knowledge Integration for Music Audio Analysis (22082)
- Developmental Machine Learning: From Human Learning to Machines and Back (22422)
- Differential Equations and Continuous-Time Deep Learning (22332)
- Efficient and Equitable Natural Language Processing in the Age of Deep Learning (22232)
- Human-Centered Artificial Intelligence (22262)
- Machine Learning and Logical Reasoning: The New Frontier (22291)
- Machine Learning for Science: Bridging Data-Driven and Mechanistic Modelling (22382)
- Recent Advancements in Tractable Probabilistic Inference (22161)
- Towards a Unified Model of Scholarly Argumentation (22432)
- Visual Text Analytics (22191)

Cryptography, Security, Privacy

- Digital Twins for Cyber-Physical Systems Security (22171)
- Intelligent Security: Is "AI for Cybersecurity" a Blessing or a Curse (22412)
- Privacy in Speech and Language Technology (22342)
- Privacy Protection of Automated and Self-Driving Vehicles (22042)
- Security of Decentralized Financial Technologies (22421)
- Security of Machine Learning (22281)
- Symmetric Cryptography (22141)

Databases, Information Retrieval, Data Mining

- Database Indexing and Query Processing (22111)
- Knowledge Graphs and Their Role in the Knowledge Engineering of the 21st Century (22372)

Data Structures, Algorithms, Complexity

- Algebraic and Analytic Methods in Computational Complexity (22371)
- Algorithmic Aspects of Information Theory (22301)
- Computation and Reconfiguration in Low-Dimensional Topological Spaces (22062)
- Computational Social Dynamics (22452)
- Counting and Sampling: Algorithms and Complexity (22482)
- Data-Driven Combinatorial Optimisation (22431)
- Dynamic Graph Algorithms (22461)
- Estimation-of-Distribution Algorithms: Theory and Applications (22182)
- Logic and Random Discrete Structures (22061)
- Graph Embeddings: Theory Meets Practice (22132)
- Optimization at the Second Level (22441)
- Tensor Computations: Applications and Optimization (22101)
- The Constraint Satisfaction Problem: Complexity and Approximability (22201)
- Theory and Practice of SAT and Combinatorial Solving (22411)
- Theory of Randomized Optimization Heuristics (22081)
- Vertex Partitioning in Graphs: From Structure to Algorithms (22481)

Distributed Computation, Networks, Architecture, Systems

- Formal Methods and Distributed Computing: Stronger Together (22492)
- Inter-Vehicular Communication – From Edge Support to Vulnerable Road Users II (22512)
- Power and Energy-Aware Computing on Heterogeneous Systems (PEACHES) (22341)
- Towards More Flexible and Automated Communication Networks (22471)

Geometry, Image Processing, Graphics, Visualization

- 3D Morphable Models and Beyond (22121)
- Bringing Graph Databases and Network Visualization Together (22031)
- Interactive Visualization for Fostering Trust in ML (22351)
- Set Visualization and Uncertainty (22462)
- Visualization and Decision Making Design Under Uncertainty (22331)
- Visualization Empowerment: How to Teach and Learn Data Visualization (22261)

Software Technology, Programming Languages

- Educational Programming Languages and Systems (22302)
- Foundations for a New Perspective of Understanding Programming (22402)
- Model-Driven Engineering of Digital Twins (22362)
- Principles of Contract Languages (22451)
- The Human Factors Impact of Programming Error Messages (22052)
- Theories of Programming (22231)

Verification, Logic, Formal Methods, Semantics

- Finite and Algorithmic Model Theory (22051)
- New Perspectives in Symbolic Computation and Satisfiability Checking (22072)

6.1 Mobility Data Science

Organizers: Mohamed Mokbel, Mahmoud Sakr, Li Xiong, and Andreas Züfle

Seminar No. 22021

Date: January 9–14, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.1.1

Participants: Taylor Anderson, Amr Magdy, Mahmoud Sakr, Flora Salim, Maxime Schoemans, Bettina Speckmann, Marc van Kreveld, Andreas Züfle

Remote Participants: Jussara Almeida, Gennady Andrienko, Natalia V. Andrienko, Walid Aref, Eric Auquiere, Yang Cao, Sanjay Chawla, Reynold Cheng, Panos Kypros Chrysanthis, Xiqi Fei, Gabriel Ghinita, Anita Graser, Dimitrios Gunopulos, Joon-Seok Kim, Kyoung-Sook Kim, Peer Kröger, John Krumm, Johannes Lauer, Mohamed Mokbel, Mario A. Nascimento, Siva Ravada, Matthias Renz, Dimitris Sacharidis, Mohamed Sarwat, Cyrus Shahabi, Egemen Tanin, Yannis Theodoridis, Kristian Torp, Carola Wenk, Martin Werner, Song Wu, Li Xiong, Jianqiu Xu, Moustafa Youssef, Demetris Zeinalipour, Esteban Zimanyi, Dimitris Zissis

Mobility data is typically available in the form of sequences of location points with time stamps, that are generated by location tracking devices. The use of mobility data has traditionally been linked to transportation industry. Nowadays, with the availability of GPS-equipped mobile devices and other inexpensive location tracking technologies, mobility data is collected and published ubiquitously, leading to large data sets of volunteered geographic information (VGI).

In general, mobility data science is the science of transforming mobility data into (actionable) knowledge. This knowledge is critical towards solutions for traffic management, disease pandemic mitigation, micro-mobility (e.g., shared bikes and scooters), health monitoring, logistics (e.g., delivery services), to mention a few.

Despite the common goal of acquiring, managing, and generating insights from mobility data, the mobility data science community is largely fragmented, developing solutions in silos. It stems from a range of disciplines with expertise in moving object data storage and management, geographic information science, spatiotemporal data mining, ubiquitous computing, computational geometry and more. Furthermore, there is a disconnect in both industry and science between mobility data scientists and domain scientists or end users for which solutions are designed. Therefore, the goal of this Dagstuhl Seminar was to bring together and recognize the mobility data science community as an interdisciplinary research field, strengthen the definition of mobility data science, and together explore challenges and opportunities in the field. The seminar had two objectives: (1) to build a new research community of mobility data science as amalgamation of the several communities who have been looking at mobility data, and (2) to draft a research agenda for mobility data science. This Dagstuhl Seminar was the first towards these objectives. The consensus of the participants is that more events will be needed in the future to continue the community building effort.

Seminar Program

The seminar was held in the week of January 9–14, 2022. It had 47 participants specialized in different topics: data management, mobility analysis, geography, privacy, urban computing, systems, simulation, indoors, visualization, information integration, and theory. Due to COVID-19, the seminar took place in

hybrid mode, with 8 onsite, and 39 remote participants. Despite the challenge of different time zones of the participants, all sessions were attended by at least 37 participants.

The seminar program is given in Figure 6.1. In the first day, every participant gave a five-minutes self introduction, research interests, and position statement on mobility data science. The rest of the program consisted of panels, and open discussions. To work around the time zone challenge, the seminar activities were centered in the afternoon of Dagstuhl, which was still possible for the Eastern and CHN time zones. The open discussions slots were planned ad-hoc during the seminar. In particular, the slot on Tuesday was used to define what is mobility data science, or more precisely what is the scope of work of this community.

All working group and panel discussions were moderated to converge towards the seminar goals of defining a research agenda and building a community. The results are summarized in the full version of the report.

Organization and Panels

To accommodate for the hybrid mode and the time zone differences, we opted to let the participants choose to participate in one of the following three thematic working groups, each having 14–17 members and led by one of the seminar co-organizers, Mohamed Mokbel, Mahmoud Sakr, Li Xiong, and Andreas Züfle:

- **Working Group 1:** Mobility Data Acquisition and Privacy
  The scope includes the full cycle of obtaining and preparing mobility data for further processing. Examples include innovative ways of data collection, crowdsourcing, simulation, data uncertainty, data cleaning, and data visualization. It also includes innovative ways of ensuring mobile users privacy as a means of encouraging users to share their data. Results of Working Group 1 are found in Section 4 of the full version of the report.
- **Working Group 2:** Mobility Data Management and Analysis
  This includes the full data pipeline from modelling, indexing, query processing/optimization, and data analysis. Existing solutions for mobility data management were discussed and a way forward for a next generation system for mobility data management was conceived. Results of Working Group 2 are found in Section 5 of the full version of the report.
- **Working Group 3** Mobility Data Science Applications
  This working group discussed the broader impacts of mobility data science to improve understanding of human behavior, urban sustainability, improving traffic conditions, health, and situational awareness. Specific applications towards these broader impacts including map making, contact tracing, pandemic preparedness, indoor navigation, and marine transportation were discussed. Results of Working Group 3 are found in Section 6 of the full version of the report.

For each working group a dedicated panel session was organized which was attended by all seminar participants. In addition, two parallel working group sessions were held for discussions and for planning the writing of the report. The working groups presented and further discussed their results with all participants on Wednesday. Four cross-cutting panels discussed the topics of systems, funding opportunities, industry involvement, and curriculum development. All panels started with presentations of panelists as listed below, seven minutes each, where they expressed their positions concerning questions given by the panel moderator. The rest of the panel time opened the discussion to all participants.

- Mobility Data Acquisition and Privacy Panel. Moderator: Li Xiong
  Panelists: Gennady (& Natalia) Andrienko, Kyoung-Sook Kim, John Krumm, Cyrus Shahabi
- Mobility Data Management and Analysis Panel. Moderator: Mahmoud Sakr
  Panelists: Walid Aref, Panos Chrysanthis, Christian Jensen, Yannis Theodoridis
- Mobility Data Science Applications Panel. Moderator: Andreas Züfle
  Panelists: Sanjay Chawla, Flora Salim, Moustafa Youssef, Demetris Zeinalipour
- Systems Panel. Moderator: Mohamed Mokbel
  Panelists: Walid Aref, Dimitrios Gunopulos, Cyrus Shahabi, Esteban Zimányi
- Funding Opportunities Panel. Moderator: Andreas Züfle
  Panelists: Johannes Lauer, Mario Nascimento, Matthias Renz, Carola Wenk
- Industry Panel. Moderator: Mohamed Mokbel
  Panelists: John Krumm, Johannes Lauer, Siva Ravada, Mohamed Sarwat
- Curriculum Development Panel. Moderator: Mahmoud Sakr
  Panelists: Anita Graser, Marc van Kreveld, Martin Werner, Esteban Zimányi

TIme			Sunday Jan 9	Monday Jan 10	Tuesday Jan 11	Wednesday Jan 12	Thursday Jan 13	Friday Jan 14
US- Eastern	Germany - CET	China- CHN						
1:30 AM	7:30 AM	2:30 PM		Breakfast (7:30 AM - 8:45 AM)				
2:00 AM	8:00 AM	3:00 PM						
2:30 AM	8:30 AM	3:30 PM						
3:00 AM	9:00 AM	4:00 PM						
3:30 AM	9:30 AM	4:30 PM						
4:00 AM	10:00 AM	5:00 PM		Coffee	Coffee	Coffee	Coffee	Coffee
4:30 AM	10:30 AM	5:30 PM		15 x 5-min: Introductory Presentations	Open discussions	Free	Open discussions	Open discussions
5:00 AM	11:00 AM	6:00 PM						
5:30 AM	11:30 AM	6:30 PM						
6:00 AM	12:00 PM	7:00 PM		Lunch (12:15 PM - 1:30 PM)				
6:30 AM	12:30 PM	7:30 PM						
7:00 AM	1:00 PM	8:00 PM						
7:30 AM	1:30 PM	8:30 PM		Welcome	Mobility Data Management and Analysis Panel (Mahmoud)	Parallel Working Groups: Discussing Panel Outcomes	Industry Panel (Mohamed)	
8:00 AM	2:00 PM	9:00 PM		10 x 5-min: Intro Presentations				
8:30 AM	2:30 PM	9:30 PM		Coffee & cake	Coffee & cake	Coffee & cake	Coffee & cake	
9:00 AM	3:00 PM	10:00 PM		15 x 5-min: Introductory Presentations	Mobility Data Science Applications Panel (Andreas)	Presenting Working Groups Results	Curriculum Development Panel (Mahmoud)	
9:30 AM	3:30 PM	10:30 PM						
10:00 AM	4:00 PM	11:00 PM						
10:30 AM	4:30 PM	11:30 PM						
11:00 AM	5:00 PM	12:00 AM		Mobility Data Acquisition and Privacy Panel (Li)	Systems Panel (Mohamed)	Funding Opportunities (Andreas)	Parallel Working Groups: Planing the Report	
11:30 AM	5:30 PM	12:30 AM						
12:00 PM	6:00 PM	1:00 AM	Buffet Dinner: 6:00 PM	Dinner: 6:30 PM				
12:30 PM	6:30 PM	1:30 AM						
1:00 PM	7:00 PM	2:00 AM						

Fig. 6.1
Dagstuhl Seminar on Mobility Science – Program.

6.2 Mobility Data Mining: From Technical to Ethical

Organizers: Bettina Berendt, Stan Matwin, and Chiara Renso

Seminar No. 22022

Date: January 9–12, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.1.35

Remote Participants: Darren Abramson, Christine Ahrend, Bettina Berendt, Florence Chee, Thiery Chevallier, Maria Luisa Damiani, Josep Domingo-Ferrer, José Antônio Fernandes de Macedo, Sébastien Gambs, Ioannis Kontopoulos, Peter Kraus, Fen Lin, Jeanna Matthews, Stan Matwin, Fran Meissner, Anna Monreale, Francesca Pratesi, Alessandra Raffaetà, Chiara Renso, Paula Reyero-Lobo, Geoffrey Rockwell, Yannis Theodoridis, Konstantinos Tserpes, Karine Zeitouni

Mobility data is one of the fastest growing types of data, thanks to the increasing number of mobile devices approaching the population of the globe. The collection, storage and analysis of spatio-temporal data representing trajectories of moving objects is one of the topics that received major attention in the field of data analytics. The more semantic information is collected from various sources, the richer is movement data. This enriched mobility data is typically referred to as semantic trajectories. The analysis of such trajectories can produce powerful results in domains such as transportation, security, tourism, health, environment and even policy design. The recent COVID-19 outbreak shed a light on the importance of collecting mobility data for public health. However, at the same time, the more mobility data is enriched with semantics, the larger the risks of violating the privacy of users and of possible unethical uses of these data analysis results. Aspects of Computational Ethics include privacy, but they go beyond this, towards a more general vision of ethical gathering, processing, uses of data and the results of data analyses. How ethics interrelates with mobility data analysis is an emerging issue.

The objective of this Dagstuhl Seminar was therefore to start a deep interacting discussion between Mobility Data Analysis researchers and Ethics experts to link these two fields with the objective of creating the foundations of a new Mobility Data Ethics research field.

This Dagstuhl Seminar, organised by Chiara Renso, Bettina Berendt and Stan Matwin as an activity from and beyond the MASTER project [1], aimed at bringing together researchers from different disciplines from Computer Science, Mobility Analysis and Ethics to trace the path from a technical vision of mobility Analysis to an also ethics-based approach to the field.

The three-day seminar was structured into three main modules: (1) round-table presentations in which each participant presented him/her self with a question about Mobility and Ethics that represents his/her interest and an object to visualise this interest or serve as a starting point for further discussion; (2) three tutorial on "technical", "ethical" and "legal" aspects of mobility data; (3) the working groups to discuss the main topics of interest that emerged during phases (1) and (2).

As a result of the group discussions on participants' interests and the issues raised in the tutorials, we formed five main working groups:

- What is/are the trade-off(s) between data privacy and data utility?
- Mobility Data Anonymity (Can location data be really anonymous?)
- Ethics of Mobility Data: What is unique? Which guidelines?
- Mobility Data Analysis Ethics beyond the data
- Mobility Data Analysis Ethics beyond humans only: Tracking animals and moral agency

The tutorials and each of the working groups are described in a chapter of the full report. Like other Dagstuhl Seminar reports, these chapters aim at making the scientific results re-usable and extendable by others. In addition, we also want to help others profit from our experiences with the videoconferencing and other media technologies that we employed and the interaction-design choices that we made. This last chapter is a reflection also on ethical aspects of the precluded and the newly added forms of mobility of scientists (and others) in meetings during and after COVID-19.

References

1 Chiara Renso, Vania Bogorny, Konstantinos Tserpes, Stan Matwin, and José Antônio Fernandes de Macêdo. Multiple Aspect Analysis of semantic trajectories (MASTER). *Int. J. Geogr. Inf. Sci.*, 35(4):763–766, 2021.

6.3 Bringing Graph Databases and Network Visualization Together

Organizers: Karsten Klein, Juan F. Sequeda, Hsiang-Yun Wu, and Da Yan

Seminar No. 22031

Date: January 16–21, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.1.67

Participants: Michael Behrisch, Walter Didimo, Nadezhda T. Doncheva, Henry Ehlers, George Fletcher, Carsten Görg, Katja Hose, Karsten Klein, Pavel Klinov, Stephen G. Kobourov, Oliver Kohlbacher, Giuseppe Liotta, Fabrizio Montecchiani, Sebastian Müller, Catia Pesquita, Falk Schreiber, Juan F. Sequeda, Hannes Voigt, Tatiana von Landesberger, Markus Wallinger, Hsiang-Yun Wu, Da Yan

Network analytics through interactive network visualization has been essential in many research and application areas, such as bioinformatics, biomedicine, cyber security, e-commerce, social science, and software engineering. A network is often supported by graph databases with advanced query engines and indexing techniques. Graph databases have substantial contributions by academia and gained strong momentum in the industry, where the focus is on scalable systems using graph query languages that require to be learned by users. Even though the Graph Database and Network Visualization communities study the same object, a graph/network, albeit from different perspectives, they do not communicate with each other. By bringing both communities together, we aimed to initiate and foster mutual communication and joint work. The goal of this seminar was to initiate collaborative efforts, to increase the mutual awareness of each others' existing concepts and technologies, and to identify new and complementary research challenges that lead to novel scientific outcomes. We have developed the schedule for the seminar based on our experience from previous successful Dagstuhl Seminars with a balance between prepared talks, plenary discussions, and breakout groups for less structured discussions focused on a selection of highly relevant topics.

The organizers envisioned several core topics for discussion at the Dagstuhl Seminar, as outlined in the proposal:

- Integration of fundamental concepts used in the two communities
- Visual scalability and computational performance
- Visual graph query paradigm
- Responsive visualization of graph query results
- (Qualitative) Evaluation
- Domain-oriented applications

During the plenary discussions on the first day, the participants identified several more specific topics for the work in separate working groups, which lead to the following working group titles:

- Evaluation and Usefulness
- Understanding gaps and opportunities between Graph Databases and Network Visualization
- Visual querying and result visualization

Our aim was to have focused discussions on these topics in which we would be able to make significant progress during the seminar, in order to shape a position paper and to lay the foundations for subsequent collaborations. The discussions showed that there indeed is the need for a closer exchange between the communities in order to improve the mutual understanding and practical solutions, but also to identify research questions that can be tackled jointly. They however also showed the great potential in this exchange and the large interest in both communities for joint work.

We have organized the seminar during the COVID-19 pandemic. Due to various regulations and travel restrictions, only roughly half of the usual number of participants attended in person. The meeting was held purely on-site, with the exception of one participant connecting in via video conferencing. We thank Dagstuhl for equipping the seminar rooms with suitable infrastructure and for putting suitable health and safety regulations in place to create a smooth experience and safe environment for all participants.

Acknowledgments We would like to thank all participants of the seminar for their contributions and lively discussions, and the scientific directorate of Dagstuhl for providing us with the opportunity to organize this seminar. Finally, the seminar would not have been possible without the amazing support by the staff of Dagstuhl.

6.4 Privacy Protection of Automated and Self-Driving Vehicles

Organizers: Frank Kargl, Ioannis Krontiris, André Weimerskirch, and Ian Williams

Seminar No. 22042

Date: January 23–28, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.1.83

Participants: Ala'a Al-Momani, Ines Ben Jemaa, Benedikt Brecht, Michael Buchholz, Thanassis Giannetsos, Adam Henschke, Mario Hoffmann, Frank Kargl, Alexander Kiening, Ioannis Krontiris, Jason Millar, Kyriaki Noussia, Christos Papadopoulos, Jonathan Petit, Chrysi Sakellari, Yu Shang, Lauren Smith, Nataša Trkulja, Jessica Uguccioni, Bryant Walker Smith, André Weimerskirch, Ian Williams, Harald Zwingelberg

Cooperative, connected and automated mobility (CCAM) has the potential to drastically reduce accidents, travel time, and the environmental impact of road travel. To achieve their goals, connected and automated vehicles (AVs) require extensive data and machine learning algorithms for processing data received from local sensors, other cars, and road-side infrastructure. This immediately raises the question of privacy and data protection. While privacy for connected vehicles has been considered for many years, AV technology is still in its infancy and the privacy and data protection aspects for AVs are not well addressed. The capabilities of AVs pose new challenges to privacy protection, given that AVs have large sensor arrays that collect data in public spaces. Additionally, AVs capture data not only from other vehicles, but also from many other parties (i.e. pedestrians walking along a street) with very limited possibilities to offer notice and choice about data processing policies. Additionally, the driver will not necessarily be the owner of the vehicle and it may be the case that the majority of AVs are owned by fleets.

Our seminar reviewed existing technologies, standards, tools, and frameworks for protecting personal information in CCAM, investigated where such existing techniques clash with the requirements of an AV and its data processing, and identified gaps and road-blockers that need to be addressed on the way to deployment of privacy protection in AVs from a legal, technical, and ethical perspective. While we ran only a shortened online version of the originally planned seminar due to COVID pandemic limitations, we made very good progress, in particular towards identifying and structuring the challenges. Future meetings will build on the results and will discuss the different challenges in more depth, prioritize the corresponding road blockers, and push for research to overcome them.

Discussions during the seminar were organized in seven sessions with presentations from renowned experts from industry and academia, and a final discussion that collected and structured outcomes. In the concluding session, we identified four main challenges that we present in the full report alongside the talk abstracts.

- The first challenge is **ethics** and responsible behavior of companies and other actors that collect and process personal data in such systems. This goes beyond mere regulatory compliance but was seen as a promising path to complement this minimal baseline. Further discussions are required to identify ways to encourage such practices.
- Second, we discussed how **regulation** needs to evolve for future CCAM systems in order to establish a stable baseline. A challenge here will be to identify to what extent sector-specific regulation will be needed to address specifics of CCAM and if regulation of future systems is reasonable and possible.
- A third challenge is the **commercial** environment. Industry has to meet regulations and financial expectations and sometimes even conflicting goals like privacy and safety. Understanding and narrowing these trade-offs while acknowledging that industry has many such constraints that limit its flexibility requires further investigation.
- Last but not least, we see a strong progress in the privacy-enhancing **technology** (PET) as a promising path towards resolving many of the above mentioned problems. At the same time, many PETs have not been designed for the CCAM domain and might not meet its demands in data quality or latency. For this reason, we see the need to further investigate how existing PETs meet CCAM requirements or how they can be developed further to do so.

Generally speaking, there is a lack of incentives for enterprises like original equipment manufacturers (OEMs) to go beyond the legal minimum requirements to manage personal data in a privacy-respecting manner, to design privacy-preserving products, or to make the use of personal data transparent to the data subject. During our discussions one question became prominent: What could be the motivation for OEMs to do more in the field of data protection that goes beyond the bare minimum of legal compliance? Ethical and trustworthy aspects, as well as reputation and brand image could be worth investigating in answering this question. However, the field is massively interdisciplinary making it necessary to convince other involved disciplines of the value of data protection for the automotive sector.

There are several technical solutions available for protecting privacy and facilitating the privacy-by-design approach. However, the up-scaling of these solutions to larger systems and their integration with existing systems often fails because systems aspects and the related interdisciplinary issues are not taken

into account. So, further progress is needed in promoting privacy-friendly system engineering, as well as integrating PETs into complete systems, taking into consideration the special requirements of safety and trust in the automotive domain. Overall, there should be a push for joint efforts to define and deploy technologies that are superior to today's solutions and that are commercially feasible since cost and effort are split amongst many participants.

Further progress is also required for the development of best practices, methodologies, and a requirements standard similar to ISO 21434 that supports the engineering of practical privacy solutions in complex systems. This will give OEMs a proper threshold target and allow for efficient solution finding and re-use. That guidance or standard could be a layer on top of regulation, similar to how the UN ECE R155 regulation requires a Cybersecurity Management System (CSMS) for which the ISO 21434 standard defines process requirements.

6.5 Finite and Algorithmic Model Theory

Organizers: Albert Atserias, Christoph Berkholz, Kousha Etessami, and Joanna Ochremiak

Seminar No. 22051

Date: January 30–February 4, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.1.101

Participants: Isolde Adler, Albert Atserias, Christoph Berkholz, Mikolaj Bojanczyk, Andrei A. Bulatov, Thomas Colcombet, Anuj Dawar, Kousha Etessami, Diego Figueira, Erich Grädel, Martin Grohe, Sandra Kiefer, Aliaume Lopez, Jerzy Marcinkowski, Rémi Morvan, Martin Otto, David Earl Roberson, Tim Seppelt, Sebastian Siebertz, Szymon Torunczyk, Marius Tritschler, Alexandre Vigny, Igor Walukiewicz, Wei-Lin Wu, Thomas Zeume

Finite and Algorithmic Model Theory research revolves around the study of the expressive power of various logics on finite and finitely presented structures, and the connections between this and different computational models and mathematical formalisms.

The methods and tools of finite and algorithmic model theory (FAMT) have played an active role in the development of several areas of computer science. Finite or finitely representable structures are those that serve as inputs to computation, and the study of the expressive power of logical languages on such structures has led to fundamental insights in diverse areas, including database theory, computational complexity, random structures and combinatorics, verification, automata theory, proof complexity, and algorithmic game theory.

Over the past four decades FAMT has established itself as a rich research field with a strong and evolving community of researchers with a shared agenda. Much of the progress can be traced to the regular meetings of the community: the last such meeting was at Dagstuhl in 2017, and before that at Les Houches in 2012.

The principal goals of this seminar included:

1. To identify fresh challenges in FAMT arising from some of the main application areas as well as newly emerging ones.
2. To make connections between core research in FAMT and other subfields of theoretical computer science, such as the theory of combinatorial and continuous optimization algorithms, and the theory of homomorphism counts and limit structures.
3. To transfer knowledge from emerging techniques in core FAMT back to the connected subfields and application areas.
4. To strengthen the research community in FAMT, especially by integrating younger members into it.
5. To provide continuity for what has been a successful model of regular seminars for building and consolidating the productive research community in FAMT.

One of the main goals of this Dagstuhl Seminar was to capitalize on the progress and the potential impact of some of the latest developments in FAMT and related areas. Such developments include:

a) The recently established connections between symmetric models of classical computation and bounded-variable counting logics. The symmetric counterparts of classical models of computations include threshold circuits, linear and semidefinite programs, and algebraic circuits. These new results have already been used to establish new upper and lower bounds for large families of algorithms by FAMT tools.

b) The theories of homomorphism counts and limit structures in combinatorics. A recent trend of work establishes that distinguishing the structures by the number homomorphisms they admit from certain classes of patterns, or to certain classes of patterns, is a fruitful alternative to distinguishing them by the logic formulas they satisfy.

c) Enumeration and counting methods including their use in database query processing, among others. One of the goals of this line of research is to understand and classify the logical queries for which it is possible to compute a compact representation of the output from which the query results can be obtained, efficiently, and on demand.

Organization and Activities

The organizers developed a schedule consisting of a number of invited survey talks, a number of talks focused on regular contributions proposed by participants, and an open problem session.

The seminar took place in person at Dagstuhl, with essentially all talks (except one) delivered by speakers attending Dagstuhl in person. The timing of the seminar coincided with the height of the Covid Omicron wave in Germany and Europe. This resulted

in a number of late covid-related cancellations. Some talks, including invited talks, had to be cancelled, and the total number of participants (24) was fewer than originally planned.

Outcomes

Despite the many organizational challenges presented by the covid surge around the time of the workshop, the seminar was highly stimulating and surprisingly successful for those who were able to attend, and achieved many of our goals. (And thankfully there were no cases of covid during the seminar among those who did attend in person.)

The final program included invited tutorial talks by Martin Grohe on homomorphism counts (delivered online, to an in person audience at the workshop, due to a last minute cancellation for Grohe caused by COVID), David Roberson I and II on quantum isomorphism and its connection to homomorphism counting. These and other related talks at the seminar highlighted the exciting ongoing work aimed at delineating the power of homorphism counting on various classes of graphs and structures, with surprising connections to other areas of mathematics.

Another invited talk was by Albert Atserias on symmetric computation and descriptive complexity (replacing a last minute cancellation), highlighted the exciting recent developments at the intersection of FAMT and combinatorial optimization. Another invited talk on query enumeration also had to be cancelled due to COVID. The full schedule including all the other talks of the Seminar can be found in the adjoined table.

	Monday	Tuesday	Wednesday	Thursday	Friday
09:00--09:15	Welcome				
09:15--09:30	Intros		Group photo		
09:30--10:00	Martin Grohe	Albert Atserias	Anuj Dawar	Mikolaj Bojanczyk	Christoph Berkholz
10:00--10:30	Martin Grohe	Albert Atserias	Anuj Dawar	Mikolaj Bojanczyk	Wrap-up
10:30--11:00	*Coffee break*	*Coffee break*	*Coffee break*	*Coffee break*	*Coffee break*
11:00--12:00	David Roberson I	David Roberson II	Erich Graedel	Thomas Colcombet	
12:00--14:00	*Lunch Break*	*Lunch Break*	*Lunch Break*	*Lunch Break*	*Lunch*
14:00--14:30	Andrei Bulatov	Szymon Torunczyk	EXCURSION/OUTING	Marius Tritschler	
14:30--15:00	Andrei Bulatov	Szymon Torunczyk		Jerzy Marcinkowski	
15:00--15:30	Thomas Zeume			Sandra Kiefer	
15:30--16:00	*Coffee/Cake*	*Coffee/Cake*	*Coffee/Cake*	*Coffee/Cake*	
16:00--16:30		Isolde Adler		Wei-Lin Wu	
16:30--17:00		Sebastian Siebertz		Open problems	
17:00--17:30		Sebastian Siebertz		Open problems	
	Dinner	*Dinner*	*Dinner*	*Dinner*	

One of the traditions of the series of workshops on Finite and Algorithmic Model Theory is to have a session in which some of the attendants present open problems and directions for further research. In this occasion, an hour of the afternoon of Thursday was devoted to such a session. A couple of days earlier we made a public call for presentations of open problems. Volunteers would write down their name on an easel pad. By Thursday, three volunteers came forward who gave 10 minute presentations (aprox.) of their proposals.

First, Erich Graedel presented an open problem on the topic of his earlier talk on Wednesday. Shortly put, the open problem asks to develop a proof theory for the emerging field of semiring semantics for logic formulas. The motivation comes from its potential applications in database theory and game analysis. Second, Isolde Adler presented an open problem on a topic covered in an earlier talk by Torunczyk. In brief, the open problem asks to study the relationships between the various notions of model-theoretic stability for classes of hypergraphs and more general relational structures. Third, Kousha Etessami gave a short talk on his recent work on applications of Tarski's Fixed-Point Theorem to economic game theory. In a nutshell, the question asks how much faster can one compute an arbitrary fixed-point of a monotone operator on a grid lattice than it takes to compute the least or greatest fixed-point by the standard iteration method. A detailed exposition of this open problem and its motivations can be found in a later section of this report. Finally, during his talk on Tuesday, Albert Atserias announced an open problem related to the optimal hardness of approximating the minimum vertex cover on graphs. Since it was presented earlier during a talk, this open problem was not presented during the session. A detailed description of this open problem appears also in a section of the full report.

Overall, the organizers regard the seminar as a resounding success despite the difficult circumstances, and judging by the very positive feedback from participants, they agreed. We look forward to the next meeting of the FAMT community, hopefully within a few years, whether at Dagstuhl or elsewhere.

The organizers are grateful to the Scientific Directorate of the Center for its support of this workshop and the staff of Schloss Dagstuhl for their organisation of our stay (including regular covid testing) and their hospitality, despite the many challenges posed by covid.

6.6 The Human Factors Impact of Programming Error Messages

Organizers: Brett A. Becker, Paul Denny, Janet Siegmund, and Andreas Stefik

Seminar No. 22052

Date: January 30–February 2, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.1.119

Participants: Annabelle Bergum, Joe Dillane, Kerstin Haring, Elisa Madeleine Hartmann, Felienne Hermans, Tobias Kohn, Jens Mönig, Raymond Pettit, Ma. Mercedes T. Rodrigo, Eddie Antonio Santos, Janet Siegmund

Remote Participants: Brett A. Becker, Dennis Bouvier, Neil Brown, Paul Denny, Kathi Fisler, Ioannis Karvelas, Amy Ko, Michael Kölling, Shriram Krishnamurthi, Linda McIver, Jan Pearce, Prajish Prasad, James Prather, Seán Russell, Andreas Stefik, Toni Taipalus, Jan Vahrenhold

Programming error messages (commonly called compiler error messages) pose challenges to programmers – from novices to professionals – with evidence dating from the 1960s to present day. In this seminar, we explored the nature of these challenges and particularly why they remain largely unaddressed. We further investigated the specific challenges that different users including children, non-native English speakers, and those of varying ability experience when faced with programming error messages. Finally, we sought to identify the most promising avenues to assess the effectiveness of error messages, how to improve them with large, demonstrated effect, and how to produce appropriate messages for different users with different needs. To this end, we assembled experts from many sub-disciplines of Computer Science, including Programming Languages, HCI, Computer Science Education, and Software Engineering as well as Learning Sciences. Due to travel restrictions imposed by the COVID-19 pandemic, we ran the seminar in a hybrid format, with 11 on-site participants collaborating with 17 remote participants. We formulated a schedule for the seminar that sought to maximise outcomes for all participants.

By combining the expertise of these different disciplines, we could identify gaps in knowledge and high-priority areas to build a basis for future work. This is the starting point for a long-lasting contribution to the field. By uniting communities that have to date been working largely in isolation, we sought to gather appropriate and useful data, broaden perspectives, build consensus among diverse stakeholders, and begin cross-community efforts working in unison going forward.

During the seminar, we had sessions focusing on existing literature and practice, including experience reports from language maintainers, expert users, researchers, and educators, to develop a shared understanding of the evidence that exists with regard to effective programming error messages. This included the group working together to synthesise existing data sets, which we view as an important exercise given that large corpora of errors and error message data are generally not openly available. This can make it difficult for researchers to answer seemingly simple questions such as: "What are the most frequent error messages encountered by [students, professionals, blind, non-native English speakers] in language x?" or, "What are evidence-based examples of effective and/or ineffective error messages?". As a result, we identified several fruitful avenues in the form of cross-discipline collaboration, data sharing opportunities, and improved focus on what steps are needed to improve efforts to answer these questions.

One of the key objectives of the seminar was to identify areas for immediate research. Several problems have been identified, including but not limited to:

1. linking error messages with the context of the problems and programs being worked on when error messages are generated;

2. understanding the metacognitive aspects involved in the process of interpreting and reacting to error messages;
3. identifying what error messages arise from inconsistent conceptions of how a program runs or how it is structured;
4. the effects of error message presentation on error message interpretation (e.g., visual queues, structured error messages, etc.);
5. error message classification schemes;
6. metrics to measure and assess error messages;
7. identifying when error messages are "wrong", what occurs when they steer the programmer in the wrong direction, and how this can be avoided;
8. determining the design factors of programming error messages that differ across demographic groups (e.g., expertise level, disabilities, native/natural language, etc.).

Another key objective was to establish new cross-community efforts to improve programming error messages in practice, leveraging the strengths of each community. Participants discussed open research problems and identified those of high priority and interest including those listed above. Several interdisciplinary research objectives have been established, and seeds were sown to form teams to collaboratively address these research questions.

6.7 Logic and Random Discrete Structures

Organizers: Erich Grädel, Phokion G. Kolaitis, and Marc Noy

Seminar No. 22061

Date: February 6–11, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.2.1

Participants: Aida Abiad, Isolde Adler, Albert Atserias, Manuel Bodirsky, Andrei A. Bulatov, Anuj Dawar, Valentin Féray, Erich Grädel, Benny Kimelfeld, Peter Lindner, Tobias Müller, Matthias Naaf, Marc Noy, Oleg Pikhurko, Peter Rossmanith, Dan Suciu, Lidia Tendera, Teun Verstraaten

Remote Participants: Andreas R. Blass, Valentin Goranko, Batya Kenig, Lefteris M. Kirousis, Phokion G. Kolaitis, Daniel Král', Jerzy Marcinkowski, Patrice Ossona de Mendez, Sudeepa Roy, Fiona Skerman, Christoph Standke, Szymon Torunczyk, Maksim Zhukovskii

Topic and Goals of the Seminar The analysis of large random discrete structures, such as trees, graphs, or permutations, is a focus of research in contemporary discrete mathematics. Logic provides a useful and powerful formalism for expressing and classifying discrete structures; moreover, it is intimately linked to the study of algorithms, computational complexity, and structural graph theory. Over the past several decades, researchers have studied random discrete structures from a logical perspective. The first significant result in this direction was the zero-one law for first-order logic under the uniform measure; this seminal result, was followed by the discovery of 'logical limit laws' or 'convergence laws' for several different models of random discrete structures and for various logics of significance in computer science. In more recent years, a renewed impetus has emerged for research activity on random discrete structures from a logical perspective. This is in part due to the availability of new methods and techniques, including asymptotic enumeration, discrete harmonic analysis, an extension of Gowers norms, and limit structures. Exciting new results on random geometric graphs, graphs on surfaces, classes of sparse graphs, graph limits, and flag algebras have been established. On the computer science side, there has been a systematic exploration of probabilistic databases, which has brought together databases, logic, and random structures. The main aim of this seminar has been to bring together some of the foremost experts from these different fields, as well as junior researchers who may become motivated to work deeper in the frontier of logic and random structures. In addition to making tangible progress on some of the currently outstanding open problems in this area, we wanted to establish new connections between (classical) random discrete structures, flag algebras, and sparse graph limits, both in terms of identifying new research questions and embarking on new collaborations, as well as fruitful interaction between foundational research and different application areas, including probabilistic databases.

Organisation and Activities Despite the restrictions and problems caused by corona pandemic, the seminar had originally been intended as a non-hybrid event with all participants on site. At the end, however, this turned out to be infeasible; as a result, two of the invited survey talks and a number of the contributed talks had to be given remotely via Zoom.

The organisers created a schedule consisting of four invited one-hour survey talks, and more focussed regular contributions proposed by the participants. The survey talks were given by

- Albert Atserias on certifying the chromatic number of a random graph;
- Fiona Skerman on the inference of underlying community structures in partially observed graphs;
- Dan Suciu on probabilistic databases;
- Patrice Ossona de Mendez on limits of graphs.

The talks of Fiona Skerman and Patrice Ossona de Mendez were given over Zoom.

In addition, there were 18 contributed talks, 11 of which were given on site, and 7 remotely via Zoom.

Overall, the organisers regard the seminar to have been a very successful scientific event. There was a general view shared by all participants that the community working on logic and random structures is in excellent shape, with interesting new developments and exciting results in many different directions. The participants clearly expressed the wish to a have a future meeting of this community, be in in Dagstuhl or elsewhere, within the next two to three years.

The organisers are grateful to the Scientific Directorate and to the staff of the Center for their support of this seminar.

6.8 Computation and Reconfiguration in Low-Dimensional Topological Spaces

Organizers: Maike Buchin, Arnaud de Mesmay, Anna Lubiw, and Saul Schleimer

Seminar No. 22062

Date: February 6–11, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.2.17

Participants: Florestan Brunck, Kevin Buchin, Maike Buchin, Benjamin Burton, Jean Cardinal, Éric Colin de Verdière, Arnaud de Mesmay, Linda Kleist, Boris Klemz, Irina Kostitsyna, Francis Lazarus, Maarten Löffler, Alexander Neuhaus, Tim Ophelders, Alexander Wolff

Remote Participants: Mark Bell, Hsien-Chih Chang, Vincent Delecroix, Anne Driemel, Nathan Dunfield, William Evans, Brittany Terese Fasy, Fabrizio Frati, Niloufar Fuladi, Elise Goujard, Luke Jeffreys, Anna Lubiw, Torsten Mütze, Hugo Parlier, Lionel Pournin, Jessica S. Purcell, Saul Schleimer, Lena Schlipf, Eric Sedgwick, Rodrigo I. Silveira, Jonathan Spreer, Stephan Tillmann, Uli Wagner, Carola Wenk, Erin Moriarty Wolf Chambers

This seminar was proposed as a followup to the Dagstuhl Seminars 17072: "Applications of Topology to the Analysis of 1-Dimensional Objects" and 19352: "Computation in Low-Dimensional Geometry and Topology". The goal of these seminars was to bring together researchers from different communities who are working on low-dimensional topological spaces (curves, embedded graphs, knots, surfaces, three-manifolds), in order to foster collaborations and synergies. Indeed, while the mathematical study of these objects has a rich and old history, the study of their algorithmic properties is still in its infancy, and new questions and problems keep coming from theoretical computer science and more applied fields, yielding a fresh and renewed perspective on computation in topological spaces.

The success of previous seminars demonstrated that research in low-dimensional topology is very active and fruitful, and also that there was a strong demand for a new seminar gathering researchers from the various involved communities, namely geometric topology and knot theory, computational geometry and topology, all the way to graph drawing and trajectory analysis.

For this iteration we placed a particular emphasis on topics related to geometric and topological **reconfiguration**: How can one structure be changed into another? How far apart are two structures? Such questions lie at the heart of various geometric problems such as computing the Fréchet distance as a way to quantify curve similarity, or morphing between two versions of a common graph. In many cases, the combinatorics and the geometry of a reconfiguration space also emerged as important objects of study: examples include associahedra, the flip graphs of triangulations, and the curve complexes in geometric topology.

The seminar started with four overview talks given by researchers in geometric topology, computational geometry, topological dynamics, and graph drawing to motivate and propose open problems that would fit the diverse backgrounds of participants and the specific focus on reconfiguration chosen for this year's workshop. This was followed by an open problem session where we gathered fifteen open problems, some of which were circulated in advance of the meeting. The remainder of the week was spent actively working on solving these problems in small groups.

The Covid pandemic prevented many participants from attending the seminar physically, and the entirety of the seminar took place in a hybrid setting, with most working groups featuring both online and physical participants. In order to coordinate the progress, we used Coauthor, a tool designed by Erik Demaine (MIT), which greatly facilitated the collaborations, and also allowed participants to have a record of the work when the seminar concluded. We also held two daily progress report meetings, allowing people to share progress and allow people to switch groups. In addition to the traditional hike, a virtual social meeting was held on Gather.town to foster interactions between the online and the physical participants.

We now briefly describe the problems that have been worked on, with a more in-depth survey of the problems and the progress being done being featured farther down in this Dagstuhl Report. Some more open problems that have been proposed but not worked on are also listed at the end of the document.

Two groups worked on questions pertaining to reconfiguring curves in the plane and on surfaces. The group 4.1 investigated problems inspired by nonograms, where one aims at introducing switches at intersections of curves in the plane to remove so-called popular faces. The group 4.5 looked at the reconfiguration graph obtained under the action of local moves on minimal closed (multi-)curves on surfaces, and whether such multi-curves could be realized as the set of geodesics of some hyperbolic metric on the surface.

A different flavor of surfaces was studied by the group 4.4, who investigated how square-tiled surfaces could be transformed under the action of shears of cylinder blocks.

The working group 4.2 studied the longstanding problem of the computational complexity of evaluation the rotation distance between elimination trees in graphs. A different flip graph, namely the one of order-k Delaunay triangulations was the topic of study of group 4.7.

Finally, two groups worked on motion of discrete objects in different contexts. The group 4.3 initiated a generalization of the classical theory of morphings of planar graph when one allows the morph to go through a third dimension. The group 4.6 investigated Turning machines, which is a simple model of molecular robot aiming to fold into specific shapes.

All in all, the seminar fostered a highly collaborative research environment by allowing researchers from very diverse backgrounds to work together on precise problems. While the hybrid setting proved to be a significant challenge, the quality of the equipment at Dagstuhl and the online tools that were used provided a practical way for all the participants to interact and to make progress on problems related to reconfiguration in geometric and topological settings.

6.9 New Perspectives in Symbolic Computation and Satisfiability Checking

Organizers: Erika Abraham, James H. Davenport, Matthew England, and Alberto Griggio

Seminar No. 22072

Date: February 13–18, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.2.67

Participants: Erika Abraham, Christopher W. Brown, Alessandro Cimatti, James H. Davenport, Tereso del Rio, Matthew England, Jürgen Gerhard, Alberto Griggio, Tudor Jebelean, Konstantin Korovin, Gereon Kremer, Corin Lee, Sergio Mover, Jasper Nalbach, Thomas Powell, Stefan Ratschan, Ali K. Uncu, Baptiste Vergain

Remote Participants: Haniel Barbosa, Anna Maria Bigatti, Russell Bradford, Martin Brain, Curtis Bright, Martin Bromberger, Changbo Chen, Isabela Dramnesc, Bruno Dutertre, Madalina Erascu, Pascal Fontaine, Vijay Ganesh, Ahmed Irfan, Manuel Kauers, Daniela Kaufmann, Ilias S. Kotsireas, Enrico Lipparini, Chenqi Mou, Mizuhito Ogawa, AmirHosein Sadeghimanesh, Martina Seidl

Introduction

Symbolic Computation refers to algorithms for computers to perform symbolic mathematics, usually implemented in Computer Algebra Systems (CASs). Satisfiability Checking refers to algorithms to efficiently check the satisfiability of a logical statement, developed originally for the Boolean domain and implemented in SAT solvers, but now extended to a wide variety of different theories in satisfiability modulo theories (SMT) solvers. This Dagstuhl Seminar is on Symbolic Computation and Satisfiability Checking, with the emphasis on the "and" to indicate the scope is strictly work of interest to both communities.

Traditionally, the two communities have been largely disjoint and unaware of the achievements of one another, despite there being strong reasons for them to discuss and collaborate, since they share many central interests. Many of the theories tackled by SMT have been traditionally studied within Symbolic Computation; while in the opposite direction, the integration of SAT solvers into computer algebra systems can allow more powerful logical reasoning and inspire new algorithmic approaches in computer algebra.

Recent History

The first global meeting dedicated to both symbolic computation and satisfiability checking was Dagstuhl Seminar 15471 (Symbolic Computation and Satisfiability Checking) [1] which took place in November 2015. This was followed soon after by EU Horizon 2020 Grant 712689 which ran from 2016–2018. The aim of that project was to bridge the gap between the communities to produce individuals who can combine the knowledge and techniques of both fields to resolve problems currently beyond the scope of either [2]. The project funded new collaborations, new tool integrations, proposals on extensions to the SMT-LIB language standards, new collections of benchmarks, two summer schools (in 2017 and 2018) and the SC-Square Workshop Series.

The Workshop Series (http://www.sc-square.org/workshops.html) has taken place annually for six years, with two further editions already planned:

2016 Timişoara, Romainia (as part of SYNASC 2016).

2017 Kaiserslautern, Germany (alongside ISSAC 2017).

2018 Oxford, UK (as part of FLoC 2018).

2019 Bern, Switzlerland (as part of SIAM AG19)

2020 Paris, France (online) (alongside IJCAR 2020)

2021 Texas, USA (online) (as part of SIAM AG21)

2022 Haifa, Israel (as part of FLoC 2022)

2023 Tromsø Norway (alongside ISSAC 2023)

It takes place as part of, or alongside, established conferences (alternating between computational algebra and logic). Each year there are two chairs, one from each community.

In 2020 a special issue of the Journal of Symbolic Computation was published, on the theme of SC-Square [3]. A further special issue is in development.

Motivation for new Seminar

The seminar call defined its scope with these research questions.

Decision Procedures: How to efficiently leverage CAS for SMT over hard arithmetical theories? How to exploit conflict-driven learning and non-chronological backtracking in symbolic computation algorithms? How can CAS and SMT be combined to reason about bit-precise machine (i.e., floating point) arithmetic?

Abstraction and Linearization: How can abstraction techniques commonly adopted in SMT be exploited in symbolic computation? How to leverage techniques in CASs for iterative abstraction refinement in SMT?

Optimization: Can SMT and symbolic computation be combined for successfully attacking non-linear optimization problems? Can new optimization techniques be leveraged for heuristic choices in solvers?

Machine Learning: What are the common challenges and opportunities on the use of Machine Learning (ML) for heuristic choices in algorithms? How best to define problem features for classic ML? How best to encode formulae for deep ML? How to develop good datasets for ML? Tool development: How to share data structures, low-level libraries, input formats and interaction pipelines for more effective development of robust, mature and interoperable symbolic reasoning tools?

Application Problem Encoding: How best encode high-level application problems to be more amenable to symbolic reasoning? How to provide more expressive problem definition languages which can still be handled efficiently? How to automate problem encoding?

Seminar Overview

The seminar was organised into eight session by broad topic (with some exceptions to allow for online participants). We invited three extended tutorials on key topics of interest to the seminar: Ahmed Irfan spoke on the incremental linearization techniques developed for MathSAT to tackle non-linear problems, including ones involving transcendental functions; Haniel Barbosa and Gereon Kremer described the new work on proof certificates in CVC5, and the possibilities for extensions into non-linear real arithmetic; and Curtis Bright spoke on isomorphism free exhaustive generation techniques which used a combination of computer algebra and SAT solvers. Other talks were contributed by seminar participants.

Upcoming Development

Erika Ábrahám, Chris Brown, James Davenport, Pascal Fontaine and Thomas Sturm are the joint editors of a Journal of Symbolic Computation Special Issue on the topic of "Symbolic Computation and Satisfiability Checking". Contributions coming out of this workshop would be especially welcomed. The timeline is given below.

31 March Submissions Open

31 August Submissions Close
(early notification to `abraham@cs.rwth-aachen.de` is welcomed.)

31 December Authors notified

3–6 months Articles published
Special issues are now "virtual" and so the articles appear online as ready.

Acknowledgements

The present seminar was originally scheduled for November 2020 but was re-organised for February 2022 following the COVID19 pandemic. It was held in a hybrid format, with participants both in person and online. The organisers thank all participants for their contribution, and extend a special thanks to the team at Schloss Dagstuhl for their excellent organisation and support, particularly in the difficult circumstance which the pandemic imposed.

References

1 E. Ábrahám and P. Fontaine and T. Sturm and D. Wang. *Symbolic Computation and Satisfiability Checking (Dagstuhl Seminar 15471)*. Dagstuhl Reports 5(11):71–89. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2016. URL: https://doi.org/10.4230/DagRep.5.11.71

2 E. Ábrahám, J. Abbott, B. Becker, A.M. Bigatti, M. Brain, B. Buchberger, A. Cimatti, J.H. Davenport, M. England, P. Fontaine, S. Forrest, A. Griggio, D. Kroening, W.M. Seiler, and T. Sturm. *SC2: Satisfiability checking meets symbolic computation*. In M. Kohlhase, M. Johansson, B. Miller, L. de Moura, and F. Tompa, editors, Intelligent Computer Mathematics: Proceedings CICM 2016, volume 9791 of Lecture Notes in Computer Science, pages 28–43. Springer International Publishing, 2016. URL: https://doi.org/10.1007/978-3-319-42547-4_3

3 J.H. Davenport, M. England, A. Griggio, T. Sturm, and C. Tinelli. *Symbolic computation and satisfiability checking: Editorial.* Journal of Symbolic Computation, 100:1–10, 2020. URL: https://doi.org/10.1016/j.jsc.2019.07.017

6.10 Theory of Randomized Optimization Heuristics

Organizers: Anne Auger, Carlos M. Fonseca, Tobias Friedrich, and Johannes Lengler

Seminar No. 22081

Date: February 20–25, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.2.87

Participants: Mickaël Binois, Alexandre Chotard, Duc-Cuong Dang, Benjamin Doerr, Carola Doerr, Carlos M. Fonseca, Armand Gissler, Tobias Glasmachers, Andreia P. Guerreiro, Nikolaus Hansen, Ekhine Irurozki, Martin S. Krejca, Johannes Lengler, Xiaoyue Li, Peter Richtarik, Günter Rudolph, Sebastian U. Stich, Dirk Sudholt, Andrew M. Sutton, Ricardo Takahashi, Elizabeth Wanner

This seminar is part of a biennial seminar series. This year, we focused on connections between classical topics of the community, such as Evolutionary Algorithms and Evolutionary Strategies (EA, ES), Estimation-of-Distribution Algorithms (EDA) and Evolutionary Multi-Objective Optimization (EMO), and related fields like Stochastic Gradient Descent (SGD) and Bayesian Optimization (BO). The mixture proved to be extremely successful. Already the first talk turned into a two hours long, vivid and productive plenary discussion. Participants like Peter Richtarik and Sebastian Stich brought valuable new perspectives from the SGD community, and Mickaël Binois contributed the BO perspective. This yielded some new approaches to long-standing open problems, specifically for a convergence proof of the CMA-ES algorithm on quadratic functions.

Another interesting and fruitful aspect of the seminar was a shift of perspective to search spaces that are under-represented in the community. Traditionally, the search spaces are product spaces, either discrete (especially the n-dimensional hypercube), or continuous (d-dimensional Euclidean space). This year we had some intense discussions in plenum and in working groups on other search spaces, triggered especially by Ekhine Irurozki's presentation on permutation spaces.

Naturally, a big part of the seminar was also devoted to classical topics of the community. Highlights included talks by Benjamin Doerr on the first runtime result for the Non-Dominated Sorting Genetic Algorithm (NSGA-II) and by Tobias Glasmachers on Convergence Analysis of the Hessian Estimation Evolution Strategy (HE-ES). The latter is the first convergence proof for a covariance matrix algorithm that does not truncate the condition number of the estimated covariance matrix. Some interesting new topics were also identified in traditional fields, such as whether we can understand better in which situations adapativity is necessary for efficient optimization by considering k-adaptive query complexity of optimization benchmarks.

Overall, as organizers we were extremely happy with the mix of core community members and researchers from related fields. The connections with the latter were close enough that scientific discussions could (also) happen on technical levels, which is particularly useful since some low-hanging fruits are available from such interchanges. Importantly, the exchange happened between people who would probably not have met each other outside of the Dagstuhl Seminar.

The seminar took place during the peak of the Omicron wave of Covid19, which made planning very difficult. The key step during preparation phase was a survey among the participants a few weeks before the seminar. We asked how likely it was that they could participate in person, and under which circumstances they would prefer which format (in-person or hybrid). The participants signalled us very clearly that they wanted this event to happen, and that they wanted it to happen in person. We want to thank all participants for their support! Other seminars in the week before and after ours had to be cancelled altogether, and this might also have happened to our seminar if not for the determination of our participants.

The seminar was smaller than previous versions, due to corona regulations. Moreover, some participants had to cancel at the last moment because they were corona-positive, or because they had no reliable child care. Especially the latter point can be frustrating, and we hope that Dagstuhl will be able to resume their support for on-site child care in the future. On the positive side, the intensity of the seminar more than made up for the smaller size, and might even have been due to the smaller number of participants.

Finally, we want to thank Dagstuhl for their great support, both financially and to their great staff. We could always feel that it was their top priority to help us, and we are greatly indebted for the support!

The organizers,

Anne Auger, Carlos M. Fonseca, Tobias Friedrich, Johannes Lengler

6.11 Deep Learning and Knowledge Integration for Music Audio Analysis

Organizers: Rachel Bittner, Meinard Müller, and Juhan Nam

Seminar No. 22082

Date: February 20–25, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.2.103

Participants: Stefan Balke, Rachel Bittner, Alice Cohen-Hadria, Simon Dixon, Simon Durand, Sebastian Ewert, Magdalena Fuentes, Dasaem Jeong, Michael Krause, Cynthia Liem, Gabriel Meseguer Brocal, Meinard Müller, Juhan Nam, Yigitcan Özer, Geoffroy Peeters, Gaël Richard, Umut Simsekli, Daniel Stoller, Bob Sturm, Gül Varol, Yu Wang, Christof Weiß

This executive summary gives an overview of our discussions on the integration of musical knowledge in deep learning approaches while summarizing the main topics covered in this seminar. We also describe the seminar's group composition, the overall organization, and the seminar's activities. Finally, we reflect on the most important aspects of this seminar and conclude with future implications and acknowledgments.

Overview

Music is a ubiquitous and vital part of our lives. Thanks to the proliferation of digital music services, we have access to music nearly anytime and anywhere, and we interact with music in a variety of ways, both as listeners and active participants. As a result, music has become one of the most popular categories of multimedia content. In general terms, music processing research aims to contribute concepts, models, and algorithms that extend our capabilities of accessing, analyzing, understanding, and creating music. In particular, the development of computational tools that allow users to find, organize, analyze, generate, and interact with music has become central to the research field known as Music Information Retrieval (MIR). Given the complexity and diversity of music, research has to account for various aspects such as the genre, instrumentation, musical form, melodic and harmonic properties, dynamics, tempo, rhythm, timbre, and so on.

As in general multimedia processing, many of the recent advances in MIR have been driven by techniques based on deep learning (DL). For example, DL-based techniques have led to significant improvements for numerous MIR tasks including music source separation, music transcription, chord recognition, melody extraction, beat tracking, tempo estimation, and lyrics alignment. In particular, major improvements could be achieved for specific music scenarios where sufficient training data is available. A particular strength of DL-based approaches is their capability to extract complex features directly from raw audio data, which can then be used for making predictions based on hidden structures and relations. Furthermore, powerful software packages allow for easily designing, implementing, and experimenting with machine learning models based on deep neural networks (DNNs).

However, DL-based approaches also come at a cost, being a data-hungry and computing-intensive technology. Furthermore, the design of suitable network architectures (including the adaption of hyper-parameters and optimization strategies) can be cumbersome and time-consuming – a process that is commonly seen more as an art rather than a science. Finally, the behavior of DL-based systems is often hard to understand; the trained models may capture information that is not directly related to the core problem. These general properties of DL-based approaches can also be observed when analyzing and processing music, which spans an enormous range of forms and styles – not to speak of the many ways music may be generated and represented. While one aims in music analysis and classification problems at capturing musically relevant aspects related to melody, harmony, rhythm, or instrumentation, data-driven approaches often capture confounding factors that may not directly relate to the target concept (e.g., recording conditions in music classification or loudness in singing voice detection).

One main advantage of classical knowledge-based engineering approaches is that they result in explainable and explicit models that can be adjusted intuitively. On the downside, such hand-engineered approaches not only require profound signal processing skills as well as domain knowledge, but also may result in highly specialized solutions that cannot be directly transferred to other problems.

As mentioned earlier, one strong advantage of deep learning is its ability to learn, rather than hand-design, features as part of a model. Nowadays, it seems that attaining state-of-the-art

solutions via machine learning depends more on the availability of large quantities of data rather than the sophistication of the approach itself. In this seminar, we critically questioned this statement in the context of concrete music analysis and processing applications. In particular, we explored existing approaches and new directions for combining recent deep learning approaches with classical model-based strategies by integrating knowledge at various stages in the processing pipeline.

There are various ways how one may integrate prior knowledge in DL-based MIR systems. First, one may exploit knowledge already at the input level by using data representations to better isolate information known to be relevant to a task and remove information known to be irrelevant (e.g., by performing vocal source separation before transcribing lyrics). Next, one may incorporate musical knowledge via the model architecture in order to force the model to use its capacity to characterize a particular aspect (e.g., limited receptive fields to prevent a model from "seeing" too much or introducing constraints that mimic DSP systems). Furthermore, the hidden representations can be conditioned to provide humans with "musically sensible control knobs" of the model (e.g., transforming an embedding space to separate out different musical instruments). Knowledge can also be exploited in the design of the output representation (e.g., structured output spaces for chord recognition that account for bass, root, and chroma) or the loss function used for optimization. During the data generation and training process, one may use musically informed data augmentations techniques to enforce certain invariances (e.g., applying pitch shifting to become invariant to musical modulations). Exploiting musical knowledge by combining deep learning techniques with ideas from classical model-based approaches was a core topic of this seminar.

The success of deep learning approaches for learning hidden structures and relations very much depends on the availability of (suitably annotated and structured) data. Therefore, as one fundamental topic, we discussed aspects of generating, collecting, accessing, representing, annotating, preprocessing, and structuring music-related data. These issues are by far not trivial. First of all, music offers a wide range of data types and formats, including text, symbolic data, audio, image, and video. For example, music can be represented as printed sheet music (image domain), encoded as MIDI or MusicXML files (symbolic domain), and played back as audio recordings (acoustic domain). Then, depending on the MIR task, one may need to deal with various types of annotations, including lyrics, chords, guitar tabs, tapping (beat, measure) positions, album covers, as well as a variety of user-generated tags and other types of metadata. To algorithmically exploit the wealth of these various types of information, one requires methods for linking semantically related data sources (e.g., songs and lyrics, sheet music and recorded performances, lead sheet and guitar tabs). Temporal alignment approaches are particularly important to obtain labels for automatic music transcription and analysis tasks. As for data accessibility, copyright issues are the main obstacle for distributing and using music collections in academic research. The generation of freely accessible music (including music composition, performance, and production) requires considerable effort, experience, time, and cost.

Besides the quantity of raw music data and its availability, another crucial issue is the input representation used as the front-end of deep neural networks. For example, log-frequency or Mel spectrograms are often used as input representations when dealing with music signals. We discussed recent research efforts where one tries to directly start with the raw waveform-based audio signal rather than relying on hand-engineered audio representations that exploit domain knowledge. In this context, we discussed how one might resolve phase shift issues by using carefully designed neural network architectures. Further recent research directions include the design of network layers to mimic common front-end transforms or incorporate differentiable filter design methods into a neural network pipeline.

Another central topic we discussed during our seminar was how to exploit musical structures via *self-supervised* and *semi-supervised learning*. Instead of relying on large amounts of labeled data, this technique exploits known variants and invariants of a dataset, using lots of *unlabeled* data. For example, without knowing the transcription of a musical piece, we know how the transcription would change if we shift the whole audio signal by some number of semitones. As another example, we can learn a notion of audio similarity by exploiting the fact that samples from a single musical audio signal are more similar than two samples drawn from different musical audio signals. We also discussed using multi-modal data to give implicit labels, such as text, image, video, and audio correspondences. On the semi-supervised learning side, representations learned in a self-supervised way can be fine-tuned to a particular task with a small amount of labeled data. In this vein, we discussed model generalization, model adaptability, active learning, few-shot learning, and human-in-the-loop systems.

Finally, we addressed topics related to the evaluation of MIR systems. In particular, we discussed the gap between loss functions typically used for optimizing deep learning pipelines and evaluation metrics designed for evaluating specific MIR tasks. In this context, we pointed out the vulnerability of standard metrics to slight variances irrelevant to the perceived output quality, expressing the need for more reliable evaluation metrics. Furthermore, we envisioned the possibility of closing the gap by designing more meaningful loss functions that may be used in the context of end-to-end learning systems.

Participants and Group Composition

In our seminar, we had 22 participants, who came from various locations around the world, including North America (2 participants from the United States), Asia (2 participants from South Korea), and Europe (18 participants from France, Germany, Netherlands, Sweden, United Kingdom). The number of participants and international constellation are remarkable considering the ongoing pandemic. (Note that many of the invited participants, particularly from overseas, were not allowed to go on business trips.) More than half of the participants (12 out of 22) came to Dagstuhl for the first time and expressed enthusiasm about the open and retreat-like atmosphere. Besides its international character, the seminar was also highly interdisciplinary. While most of the participating researchers are working in music information retrieval, we also had participants with a background in musicology, signal processing, machine learning, mathematics, computer vision, and other fields. Our seminar stimulated cross-disciplinary discussions by having experts working in technical and non-technical disciplines while highlighting opportunities for new collaborations among our attendees. Furthermore, the number of participants from the industry (6 out of 22) was relatively high, which also underlines the relevance of the seminar's topic beyond fundamental research. Most of the participants had a strong musical background, some of them even having a dual career in an engineering discipline and music. This led to numerous social activities, including playing music together. In addition to geographical locations and research disciplines, we tried to foster variety in terms of seniority levels (e.g., we had three Ph.D. students and six participants on the postdoc/junior/assistant professor level) and in terms of gender

(6 out of 22 of the participants identify as female). Besides scientific questions, we discussed in our seminar also various challenges that younger colleagues typically face when setting up their research groups and scientific curriculum at the beginning of their academic careers.

Overall Organization and Schedule

Dagstuhl Seminars have a high degree of flexibility and interactivity, allowing participants to discuss ideas and raise questions rather than presenting research results. Following this tradition, we fixed the schedule during the seminar asking for spontaneous contributions with future-oriented content, thus avoiding a conference-like atmosphere, where the focus tends to be on past research achievements. After the organizers gave an overview of the Dagstuhl concept, we started the first day with self-introductions, where all participants introduced themselves and expressed their expectations and wishes for the seminar. We then continued with short (15 to 20 minutes) stimulus talks, where specific participants addressed some critical questions related to the seminar's overall topic in a non-technical fashion. Each of these talks seamlessly moved towards an open discussion among all participants, where the respective presenters took over the role of a moderator. These discussions were well received and often lasted for more than half an hour. The first day closed with a brainstorming session on central topics covering the participants' interests while shaping the overall schedule and format for the next day. We continued having stimulus tasks interleaved with extensive discussions on the subsequent days. On the second day, we split into smaller groups, each group discussing a more specific topic in greater depth. The results and conclusions of these parallel group sessions, which lasted between 60 to 90 minutes, were then presented and discussed with the plenum. However, since the overall seminar size of 22 participants was relatively small, it turned out that the division into subgroups was not necessary. Thanks to excellent group dynamics and a fair distribution of speaking time, all participants had their say and were able to express their thoughts in the plenum while avoiding a monotonous conference-like presentation format. On the last day, we enjoyed a tutorial by Umut Simsekli on some theoretical concepts behind deep learning (a topic unanimously desired by the group). We concluded the seminar with a session we called "self-outroductions" where each participant presented their personal view on the seminar's results.

While working in technical engineering disciplines, most participants also have a strong background and interest in music. This versatility significantly impacted the seminar's atmosphere, leading to cross-disciplinary intersections and provoking discussions and resulting in intensive joint music-making during the breaks and in the evenings. One particular highlight was a concert on Thursday evening organized by Cynthia Liem and Christof Weiß, where various participant-based ensembles performed a wide variety of music, including classical music, Irish folk music, and jazz.

Conclusions and Acknowledgment

There is a growing trend toward building more interpretable deep learning systems, from the data collection and generation stage, to the input and output representations, to the model structure itself. On the other hand, classical model-based approaches bring a wealth of expertise on techniques for knowledge integration in system design. The Dagstuhl Seminar gave us the opportunity for connecting experts from classical model-based approaches, deep learning-based approaches, and related interdisciplinary fields such as music perception and human-computer interaction in order to generate discussion and spark new collaborations. The generation of novel, technically oriented scientific contributions was not the main focus of the seminar. Naturally, many of the contributions and discussions were on a conceptual level, laying the foundations for future projects and collaborations. Thus, the main impact of the seminar is likely to take place in the medium and long term. Some more immediate results, such as plans to share research data and software, also arose from the discussions. As further measurable outputs from the seminar, we expect to see several joint papers and applications for funding.

Besides the scientific aspect, the social aspect of our seminar was just as important. We had an interdisciplinary, international, and interactive group of researchers, consisting of leaders and future leaders in our field. Many of our participants were visiting Dagstuhl for the first time and enthusiastically praised the open and inspiring setting. The group dynamics were excellent, with many personal exchanges and shared activities. Some scientists expressed their appreciation for having the opportunity for prolonged discussions with researchers from neighboring research fields, which is often impossible during conference-like events. At this point, we would like to let some of the participants have their say:

- Stefan Balke (pmOne – Paderborn, DE): *"Dagstuhl is always a wonderful experience, having time to think, talk, and play music. All in a relaxed atmosphere, the seminar feels like a family meeting – especially in these times."*
- Alice Cohen-Hadria (IRCAM – Paris, FR): *"Now I feel like a part of a community."*
- Dasaem Jeong (Sogang University – Seoul, KR): *"Full of insightful discussions, music, and friends in a beautiful place."*
- Cynthia Liem (TU Delft, NL): *"Dagstuhl is the one place in the world where one effectively can have a week long unconference. More deeply talking about research and new ideas, enjoying time with academic friends, with much less distraction than one would have at home, or even in a 'regular' conference. Especially coming out of a pandemic, I am realizing this is among the most valuable things in our professional life."*
- Daniel Stoller (Spotify – Bonn, DE): *"Dagstuhl brings perspectives on the big issues."*
- Yu Wang (New York University – Brooklyn, US): *"Discussion is like music: the live version is always better."*

In conclusion, our expectations for the seminar were not only met but exceeded, in particular concerning networking and community building. We want to express our gratitude to the Dagstuhl board for giving us the opportunity to organize this seminar, the Dagstuhl office for their exceptional support in the organization process, and the entire Dagstuhl staff for their excellent service during the seminar. In particular, we want to thank Susanne Bach-Bernhard and Michael Gerke for their assistance during the preparation and organization of the seminar.

6.12 AI for the Social Good

Organizers: Claudia Clopath, Ruben De Winne, Rayid Ghani, and Tom Schaul

Seminar No. 22091

Date: February 27–March 4, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.2.134

Participants: Emran Alchikh Alnajar, Asma Atamna, Sara Beery, Claudia Clopath, Ruben De Winne, Jose Pablo Folch, Jörg Franke, Mariella Goebl, Nitusima Kataraia, Stijn Koster, Rik Linssen, Subhransu Maji, Jacopo Margutti, Corina Markodimitraki, Nele Quast, Raghu Rajan, José Ricardo Rubio Valverde, Samira Said, Tom Schaul, Ruby Sedgwick, Richard Sserunjogi, Christoph Weigl

AI and ML have made impressive progress in the last few years. Long-standing challenges like Go have fallen and the technology has entered daily use via the vision, speech or translation capabilities in billions of smartphones. The pace of research progress shows no signs of slowing down, and demand for talent is unprecedented. But as part of a wider AI for Social Good trend, this seminar wanted to contribute to ensuring that the social good does not become an afterthought in the rapid AI and ML evolution, but that society benefits as a whole. The five-day seminar brought together AI and ML researchers from various universities with representatives from NGOs based in Benin, Tanzania, Uganda, The Netherlands and globally. These NGOs all pursue various social good goals, such as improving air quality, increasing agricultural productivity with the help of technology, transforming health care, providing humanitarian support, and defeating poverty. On these topics, NGOs have rich domain knowledge, just like they have vast networks with (non-)governmental actors in developing countries. Mostly, NGOs have their finger on the pulse of the challenges that the world and especially its most vulnerable inhabitants are facing today, and will be facing tomorrow. The objective of the seminar was to look at these challenges through an AI and ML lens, to explore if and how these technologies could help NGOs to address these challenges. The motivation was also that collaborations between AI and ML researchers and NGOs could benefit both sides: on the one hand, the new techniques can help with prediction, data analysis, modelling, or decision making. On the other hand, the NGOs' domains contain many non-standard conditions, like missing data, side-effects, or multiple competing objectives, all of which are fascinating research challenges in themselves. And of course, publication impact is substantially enhanced when a method has real-world impact. The seminar facilitated the exploration of possible collaborations between AI and ML researchers and NGOs through a two-pronged approach. This approach combined high-level talks and discussions on the one hand with a hands-on hackathon on the other hand. High-level talks and discussions focused first on the central concepts and theories in AI and ML and in the NGOs' development work, before diving into specific issues such as generalisability, data pipelines, and explainability. These talks and discussions allowed all participants – in a very short time-frame – to reach a sufficient level of understanding of each other's work. This understanding was the basis to then start investigating jointly through a hackathon how AI and ML could help addressing the real-world challenges presented by the NGOs. At the start of the hackathon, an open marketplace-like setting allowed AI and ML researchers and NGOs to find the best match between technological supply and demand. When teams of researchers and NGOs were established, their initial objective was not to start coding, but to define objectives, assess scope and feasibility. The intense exchanges during the hackathon allowed NGOs with a lower AI/ML maturity increased understanding of the capabilities of AI/ML and define actions to effectively start working with AI/ML. NGOs that already had a more advanced understanding and use of AI/ML technology prior to the seminar, could take their AI maturity to the next level by trying out new ML approaches, designing and testing tailored ML models, or simply exploring new partnerships. Key to this success of the hackathon – and the seminar at large – was the presence of AI/ML experts whose respective fields of expertise could seamlessly be matched with the various needs of the various NGOs. This excellent group composition also facilitated a productive discussion about guidelines on how to do effective AI for social good collaborations in the future (e.g., by focusing on long-term partnerships, and by sequencing problem scoping before data cleaning and – only in last instance – an actual hackathon).

6.13 Tensor Computations: Applications and Optimization

Organizers: Paolo Bientinesi, David Ham, Furong Huang, Paul H. J. Kelly, and P. (Saday) Sadayappan

Seminar No. 22101

Date: March 6–11, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.3.1

Participants: Cem Bassoy, Paolo Bientinesi, Simon Bonér, Albert Choen, Jeremy Cohen, Teodoro F. Collin, Jutho Haegerman, David Ham, Paul H. J. Kelly, Thomas Koehler, Lawrance Mitchell, Christos Psarras, Norman A. Rink, P. Sadayappan, Paul Springer, Edward Stow, Volker Tresp, Bora Ucar, Carsten Uphoff, Edward Valeev, Sophia Vorderwuelbecke, Connor Ward

Remote Participants: Charisee Chiw, Nadav Cohen, Edoardo Di Napoli, Rong Ge, Johnnie Gray, Vinod Grover, Furong Huang, Katharina Kormann, Jean Kossaifi, Jiajia Li, Devin Matthews, Luke Panayi, Vivek Srikumar, Edwin Miles Stoudenmire, Richard M. Veras, Qi (Rose) Yu, Pan Zhang

Linear relationships between quantities are one of the most fundamental and pervasive phenomena in mathematics, science and computing. While matrices encode linear relationships between exactly two quantities, tensors are an abstraction representing linear relationships between multiple variables. Tensor computations therefore provide an abstract language for computations that span an enormous range of application domains, including machine learning, quantum information systems, simulations based on solving partial differential equations, computational chemistry and beyond. The tensor abstraction enriches our understanding of the structure of computations, and exposes common challenges and solutions that cut across different research communities.

While the mathematics of tensors is well-developed and extensively applied across all of these applications and beyond, there is far less commonality in the software abstractions and tools deployed to execute tensor computations. This is in stark contrast to matrix computations, where common abstractions and stable interfaces have led to widely used tools that bring high performance across diverse application domains.

This seminar explored this challenge, and made significant progress towards establishing foundations for common implementations – embodying the substantial body of knowledge on high-performance tensor computation strategies in common software libraries and domain-specific program generation tools.

The seminar began with five tutorial lectures, offered by the organisers in partnership with selected leading figures in some of the relevant communities. We began by mapping some of the diverse terminology. We then provided tutorials exposing the quantitative and qualitative diversity in how different communities use tensor computations – aiming to build a common understanding of key concepts, notations, and building blocks. We focused on the following application areas:

- Quantum physics and chemistry
- Mesh-based discretisations for solution of partial differential equations
- Machine learning.

The final tutorial reviewed the challenge of establishing unifying software tools, highlighting the enormous body of work that has been done within application areas.

The second phase of the seminar consisted of more detailed presentations from the participants. These included motivating applications, but focusing on the fundamental computational workloads, methods, and performance challenges. Building on this, we also had contributions focused on implementation – low-level performance considerations, algorithmic proposals, compiler algorithms and compiler infrastructure.

In the third phase of the seminar, we separated into three teams. One explored benchmarking and datasets, another made substantial progress on proof-of-concept implementation work to connecting the high-level Tensorly library for tensor decompositions in machine learning to lower-level tensor-vector products – achieving considerable performance advantage. Finally there was also a major and continuing effort to define a common domain-specific language and compiler representation for tensor contractions that supports both high-level optimisations and the use of high-performance low-level libraries.

This 2021 seminar built on progress made at an earlier seminar with the same title, in March 2020 – which was very heavily impacted by the coronavirus pandemic. This seminar was also affected, to a lesser extent – with a reduced number of on-site participants, partly compensated by very useful engagement with researchers joining online, albeit from distant timezones.

This seminar benefited from broader engagement with application domains – partly as a result of the work that was done on tutorials – which we hope to publish in due course. It also benefited from deeper engagement with developers of high-performance building blocks. Finally, we initiated a new and continuing effort to define a common language and a common intermediate language for code generation tools.

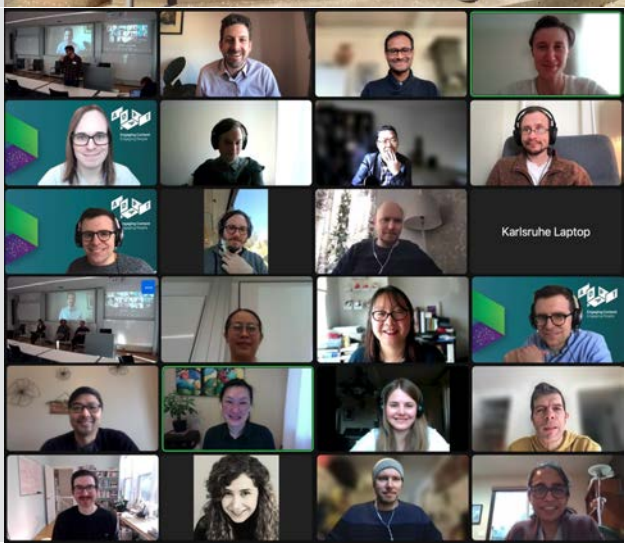
6

6.14 Computational Models of Human-Automated Vehicle Interaction

Organizers: Martin Baumann, Shamsi Tamara Iqbal, Christian P. Janssen, and Antti Oulasvirta

Seminar No. 22102

Date: March 6–11, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.3.15



Participants: Martin Baumann, Jelmer Borst, Alexandra Bremers, Duncan Brumby, Debargha Dey, Patrick Ebel, Martin Fränzle, Luisa Heinrich, Moritz Held, Jussi Jokinen, Dietrich Manstetten, Gustav Markkula, Roderick Murray-Smith, Antti Oulasvirta, Nele Rußwinkel, Shadan Sadeghian Borojeni, Hatice Sahin, Philipp Wintersberger, Fei Yan

Remote Participants: Linda Ng Boyle, Lewis Chuang, Benjamin Cowan, Birsen Donmez, Justin Edwards, Mark Eilers, Shamsi Tamara Iqbal, Christian P. Janssen, Myounghoon Jeon, Xiaobei Jiang, Wendy Ju, Tuomo Kujala, Andrew Kun, Otto Lappi, Nikolas Martelaro, Andreas Riener, Boris van Waterschoot, Yiqi Zhang

This is the executive summary of Dagstuhl 22102: Computational Models of Human-Automated Vehicle Interaction, which took place March 6–11th 2022 in Hybrid format. The executive summary first summarizes the motivation of the seminar, then gives an overview of the broad challenges that were discussed, it then presents the results of the seminar. As this is only the summary, there are a lot more details about every item and result in the full Dagstuhl Report of this seminar, which is therefore referred to.

It has been a fruitful meeting, which sparked many research ideas. We want to thank all the attendees for their attendance and all the input they generate. We hope that it is of value to the community, and we can't wait to see what other results follow in the future based on discussions that started at this seminar!

Christian Janssen, Martin Baumann, Antti Oulasvirta, and Shamsi Iqbal (organizers)

Computational Models of Human-Automated Vehicle Interaction: Summary of the field

The capabilities of automated vehicles are rapidly increasing, and are changing human interaction considerably (e.g., [4, 6, 29]). Despite this technological progress, the path to fully self-driving vehicles without any human intervention is long, and for the foreseeable future human interaction is still needed with automated vehicles (e.g., [15, 22, 29, 37, 47, 48]). The principles of human-automation interaction also guide the future outlook of the European Commission [13, 14]. Human-automated vehicle interaction can take at least two forms. One form is a partnership, in which the human and the automated vehicle both contribute in parallel to the control of the vehicle. Another form is in transitions of control, where the automated system at times takes

over full control of the vehicle, but transitions control back to the human when desired by the human, or when required due to system limitations. For both the partnership and the transition paradigm it is beneficial when the car and the human have a good model of each other's capabilities and limitations. Accurate models can make clear how tasks are distributed between the human and the machine. This helps avoid misunderstandings, or mode confusion [45], and thereby reduces the likelihood of accidents and incidents. A key tool in this regard is the use of computational (cognitive) models: computational instantiations that simulate the human thought process and/or their interaction with an automated vehicle. Computational models build on a long tradition in cognitive science (e.g., [35, 36, 44]), human factors and human-computer interaction (e.g., [10, 27, 39], neuroscience (e.g., [12, 31]), and AI and engineering (e.g., [17, 42]). By now, there are a wide set of varieties that can be applied to different domains, ranging from constrained theoretical problems to capturing real-world interaction [38]. Computational models have many benefits. They enforce a working ethic of "understanding by building" and require precision in specification ([34], see also [8, 32, 41]). Models can test the impact of changes in parameters and assumptions, which allows for wider applicability and scalability (e.g., [2, 16, 44]). More generally, this allows for testing "what if" scenarios. For human-automated vehicle interaction in particular, it allows testing of future adaptive systems that are not yet on the road. Automated driving is a domain where computational models can be applied. Three approaches have only started to scratch the surface. First, the large majority of models focus on engineering aspects (e.g., computer vision, sensing the environment, flow of traffic) that do not consider the human extensively (e.g., [7, 18, 33]). Second, models that focus on the human mostly capture manual, non-automated driving (e.g., [9, 25, 44]). Third, models about human interaction in automated vehicles are either conceptual (e.g., [20, 22]) or qualitative, and do not benefit from the full set of advantages that computational models offer. In summary, there is a disconnect between the power and capabilities that computational models offer for the domain of automated driving, and today's state-of-the-art research. This is due to a set of broad challenges that the field is facing and that need to be tackled over the next 3–10 years, which we will discuss next.

Description of the seminar topics and structure of the seminar report

The seminar topics were clustered around five broad challenges, for which we provide a brief description and example issues that were discussed and addressed. Although the challenges are presented separately, they are interconnected and were discussed in an integrated manner during the seminar. During the seminar, each challenge was discussed in a panel, with all attendees taking part in at least one panel. After each panel, the group was split up in smaller workgroups, and discussed the themes in more lengths. The summary of each panel discussion can be found in the report under the section "panel discussions". The outcomes of the workgroups can be found in the report under the section "workgroups". In addition, all attendees wrote short abstracts that summarized their individual position.

Challenge 1: How can models inform design and governmental policy?

Models are most useful if they are more than abstract, theoretical vehicles. They should not live in a vacuum, but be related to problems and issues in the real world. Therefore, we want to explicitly discuss how models can inform the design of (in-)vehicle technology, and how they can inform policy. As both of these topics can fill an entire Dagstuhl by themselves, our primary objective is to identify the most pressing issues and opportunities. For example, looking at:

- *Types of questions:* what types of questions exist at a design and policy level about human-automated vehicle interaction?
- *How to inform decisions:* How can models be used to inform design and policy decisions? What level of detail is needed here? What are examples of good practices?
- *Integration:* Integration can be considered in multiple ways. First, how can ideas from different disciplines be integrated (e.g., behavioral sciences, engineering, economics), even if they have at times opposing views (e.g., monetary gains versus accuracy and rigor)? Second, how can models become better integrated in the design and development process as tools to evaluate prototypes (instead of running empirical tests)? And third, how can models be integrated into the automation (e.g., as a user model) to broaden the automation functionality (e.g., prediction of possible driver actions, time needed to take over)?

Challenge 2: What phenomena and driving scenarios need to be captured?

The aim here is to both advance theory on human-automation interaction while also contributing to understanding realistic case studies for human-automation interaction that are faced for example by industry and governments. The following are example phenomena:

- *Transitions of control and dynamic attention:* When semi-automated vehicles transition control of the car back to the human, they require accurate estimates of a user's attention level and capability to take control (e.g., [22, 49]).
- *Mental models, machine models, mode confusion, and training and skill:* Models can be used to estimate human's understanding of the machine and vice-versa (e.g., [20]). Similarly, they might be used to estimate a human driver's skill level, and whether training is desired.
- *Shared control:* In all these scenarios, there is some form of shared control. Shared control requires a mutual understanding of human and automation. Computational models can be used to provide such understanding for the automation (e.g., [50]).

Challenge 3: What technical capabilities do computational models possess?

A second challenge has to do with the technical capabilities of the models. Although the nature of different modeling frameworks and different studies might differ [38], what do we consider the core functionality? For example, related to:

- *Compatibility:* To what degree do models need to be compatible with simulator software (e.g., to test a "virtual participant"), hardware (e.g., be able to drive a car on a test track), and other models of human thinking?
- *Adaptive nature:* Computational models aim to strike a balance between precise predictions for more static environments and being able to handle open-ended dynamic environments (like everyday traffic). How can precision be guaranteed in static and dynamic environments? How can models adapt to changing circumstances?
- *Speed of development and broader adoption:* The development of computational models requires expertise and time.

How can development speed be improved? How can communities benefit from each other's expertise?

Challenge 4: How can models benefit from advances in AI while avoiding pitfalls?

At the moment there are many developments in AI that computational models can benefit from. Three examples are advances in (1) simulator-based inference (e.g., [26]) to reason about possible future worlds (e.g., varieties of traffic environments), (2) reinforcement learning [46] and its application to robotics [30] and human driving [25], and (3) deep learning [17] and its potential to predict driver state or behavior from sensor data. At the same time, incorporation of AI techniques also comes with challenges that need to be addressed. For example:

- *Explainability:* Machine learning techniques are good at classifying data, but do not always provide insight into why classifications are made. This limits their explainability and is at odds with the objective of computational models to gain insight into human behavior. How can algorithms' explainability be improved?
- *Scalability and generalization*: How can models be made that are scalable to other domains and that are not overtrained on specific instances? How can they account for future scenarios where human behavior might be hard to predict [5]?
- *System training and corrective feedback:* if models are trained on a dataset, what is the right level of feedback to correct an incorrect action to the model? How can important new instances and examples be given more weight to update the model's understanding without biasing the impact?

Challenge 5: What insights are needed for and from empirical research?

Models are only as good to the degree as they can describe and predict phenomena in the real world. Therefore, empirical tests are an important consideration. Example considerations are:

- *Capturing behavioral change and long-term phenomena:* Many current computational models capture the results of a single experiment. However, behavior might change with more exposure to and experience with automated technology. How can such (long-term) behavior change be tested?
- *Capturing unknown future scenarios:* Many automated technologies that might benefit from computational models are not yet commercially available. How can these best be studied and connected to computational models?
- *Simulated driving versus real-world encounters:* To what degree are simulator tests representative of real-world scenarios (e.g., [19])?

Results

The seminar has generated the following results.

1. **Overview of state-of-the-art technologies, methods, and models.** The spectrum of computational modeling techniques is large [21, 24, 38]. Before and during the conference, we have discussed various methods and techniques. In particular, the full report contains a dedicated chapter called "Relevant papers for modeling human-automated vehicle interaction" in which we report a long set of papers that the community identified as being relevant to the field. We encourage scholars to take a look at it.
2. **List of grand challenges with solution paths.** We have identified five grand challenges and discussed those in detail during the panels. Our chapters on "panel discussions" report the outcomes of these discussions. Moreover, the workgroups further report the in-depth discussions that smaller groups had about these challenges. See the section "working group" of the full report. The results only start to scratch the surface of some of the grand challenges for the application of computational cognitive modeling that need to be faced within the next 3 to 10 years, and their paths to solutions. Based on discussions, groups of authors plan to work on more papers and workshops around topics that they deemed worthy of further discussion. For example, we discussed whether there are specific driving scenarios that a computational model should be able to capture, and how success might be quantified (e.g., whether these challenges should take the form of competitions, akin to DARPA's Grand Challenge for automated vehicles [11] or "Newell's test" for cognitive models [3]).
3. **Research agenda to further the field.** The full report also reports a research agenda that is intended to further the field. For each specific grand challenge, we have identified more specific areas of research that need futher exploration. We refer to the dedicated section in the full report called "Research agenda to further the field". The organizers of the seminar will also organize a dedicated journal special issue around the topic, in which further results that arose from the seminar can be reported.

References

1 Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Iqbal, S.T., and Teevan, J. (2019). Guidelines for human-AI interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1–13).

2 Anderson, J. R. (2007). How can the human mind occur in the physical universe? (Vol. 3). Oxford University Press.

3 Anderson, J. R., and Lebiere, C. (2003). The Newell test for a theory of cognition. Behavioral and brain Sciences, 26(5), 587–601.

4 Ayoub, J., Zhou, F., Bao, S., and Yang, X. J. (2019). From Manual Driving to Automated Driving: A Review of 10 Years of AutoUI. In Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 70–90).

5 Bainbridge, L. (1983). Ironies of automation. In Analysis, design and evaluation of man–machine systems (pp. 129–135). Pergamon.

6 Bengler, K., Dietmayer, K., Farber, B., Maurer, M., Stiller, C., and Winner, H. (2014). Three decades of driver assistance systems: Review and future perspectives. IEEE Intelligent Transportation Systems Magazine, 6(4), 6–22.

7 Brackstone, M., and McDonald, M. (1999). Car-following: a historical review. Transportation Research Part F: Traffic Psychology and Behaviour, 2(4), 181–196.

8 Brooks, R. A. (1991). Intelligence without representation. Artificial intelligence, 47(1–3), 139–159.

9 Brumby, D. P., Janssen, C. P., Kujala, T., and Salvucci, D. D. (2018). Computational models of user multitasking. Computational interaction design, 341–362.

10 Card, S. K., Moran, T., and Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: L. Erlbaum Associates Inc.

11 Darpa (2020) The Grand Challenge. Accessed online on July 6 2020 at https://www.darpa.mil/about-us/timeline/-grand-challenge-for-autonomous-vehicles

12 Eliasmith, C. (2013). How to build a brain: A neural architecture for biological cognition. Oxford University Press.

13 European Commission (2018, 17 May). On the road to automated mobility: An EU strategy for mobility of the future (pp. 1–17). Brussels, BE. Communication COM(2018) 283 final.

14 European Commission (2020, 19 February). Shaping Europe's digital future. Brussels (BE). Communication COM(2020) 67 final.

15 Favarò, F. M. (2020). Unsettled Issues Concerning Semi-Automated Vehicles: Safety and Human Interactions on the Road to Full Autonomy. Technical report for the SAE. Warrendale, PA: SAE International. Retrieved from https://www.sae.org/publications/technical-papers/content/epr2020001/

16 Gray, W. D. (Ed.). (2007). Integrated models of cognitive systems (Vol. 1). Oxford University Press.

17 Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep learning. Cambridge, MA: MIT press.

18 Helbing, D. (2001). Traffic and related self-driven many-particle systems. Reviews of modern physics, 73(4), 1067.

19 Hock, P., Kraus, J., Babel, F., Walch, M., Rukzio, E., and Baumann, M. (2018). How to design valid simulator studies for investigating user experience in automated driving – Review and hands-on considerations. Proceedings of the International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, 105–117. New York, NY: ACM Press

20 Janssen, C. P., Boyle, L. N., Kun, A. L., Ju, W., and Chuang, L. L. (2019). A Hidden Markov Framework to Capture Human–Machine Interaction in Automated Vehicles. International Journal of Human-Computer Interaction, 35(11), 947–955.

21 Janssen, C. P., Boyle, L. N., Ju, W., Riener, A., and Alvarez, I. (2020). Agents, environments, scenarios: A framework for examining models and simulations of human-vehicle interaction. Transportation research interdisciplinary perspectives, 8, 100214.

22 Janssen, C. P., Iqbal, S. T., Kun, A. L., and Donker, S. F. (2019). Interrupted by my car? Implications of interruption and interleaving research for automated vehicles. International Journal of Human-Computer Studies, 130, 221–233.

23 Janssen, C. P., and Kun, A. L. (2020). Automated driving: getting and keeping the human in the loop. Interactions, 27(2), 62–65.

24 Jeon, M., Zhang, Y., Jeong, H., P. Janssen, C.P., and Bao, S. (2021). Computational Modeling of Driving Behaviors: Challenges and Approaches. In 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 160–163).

25 Jokinen, J.P.P., Kujala, T., and Oulasvirta, A. (2021) Multitasking in Driving as Optimal Adaptation under Uncertainty. Human Factors, , 63(8), 1324–1341.

26 Kangasrääsiö, A., Jokinen, J. P., Oulasvirta, A., Howes, A., and Kaski, S. (2019). Parameter inference for computational cognitive models with Approximate Bayesian Computation. Cognitive Science, 43(6), e12738.

27 Kieras, D. (2012). Model-based evaluation. In: Jacko and Sears (Eds.) The Human-Computer Interaction Handbook (3rd edition), 1294–310. Taylor and Francis

28 Kun, A. L. (2018). Human-Machine Interaction for Vehicles: Review and Outlook. Foundations and Trends in Human-Computer Interaction, 11(4), 201–293.

29 Kun, A. L., Boll, S., and Schmidt, A. (2016). Shifting Gears: User Interfaces in the Age of Autonomous Vehicles. IEEE Pervasive Computing, 32–38.

30 Levine, S. (2018). Reinforcement learning and control as probabilistic inference: Tutorial and review. arXiv preprint arXiv:1805.00909.

31 Marr, D. (1982). Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. San Francisco, CA: W.A. Freeman.

32 McClelland, J. L. (2009). The place of modeling in cognitive science. Topics in Cognitive Science, 1(1), 11–38.

33 Mogelmose, A., Trivedi, M. M., and Moeslund, T. B. (2012). Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey. IEEE Transactions on Intelligent Transportation Systems, 13(4), 1484–1497.

34 Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium.In Chase (ed.) Visual Information Processing. New York: Academic Press.

35 Newell, A. (1990). Unified theories of cognition. Cambridge, MA: Harvard University Press.

36 Newell, A., and Simon, H. A. (1972). Human Problem Solving. Englewood Cliffs, NJ: Prentice-Hall

37 Noy, I. Y., Shinar, D., and Horrey, W. J. (2018). Automated driving: Safety blind spots. Safety science, 102, 68–78.

38 Oulasvirta, A. (2019). It's time to rediscover HCI models. Interactions, 26(4), 52–56.

39 Oulasvirta, A., Bi, X., Kristensson, P-O., and Howes, A., (Eds.) (2018). Computational Interaction. Oxford University Press

40 Peebles, D., and Cooper, R. P. (2015). Thirty years after Marr's vision: levels of analysis in cognitive science. Topics in cognitive science, 7(2), 187–190.

41 Pfeifer, R., and Scheier, C. (2001). Understanding intelligence. Cambridge, MA: MIT press.

42 Russell, S., and Norvig, P. (2002). Artificial intelligence: a modern approach. Uppersaddle River, NJ: Pearson

43 SAE International. (2014). J3016: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. Warrendale, PA, USA: SAE International

44 Salvucci, D. D., and Taatgen, N. A. (2011). The multitasking mind. Oxford University Press.

45 Sarter, N. B., and Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. Human factors, 37(1), 5–19.

46 Sutton, R., and Barto, A. G. (2018). Reinforcement learning: An introduction. Cambridge, MA: MIT Press

47 Walch, M., Sieber, T., Hock, P., Baumann, M., and Weber, M. (2016). Towards cooperative driving: Involving the driver in an autonomous vehicle's decision making. In Proceedings of the International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 261–268. New York, NY: ACM Press.

48 Walch, M., Mühl, K., Kraus, J., Stoll, T., Baumann, M., and Weber, M. (2017). From Car-Driver-Handovers to Cooperative Interfaces: Visions for Driver–Vehicle Interaction in Automated Driving. In G. Meixner and C. Müller (Eds.): Automotive User Interfaces: Creating Interactive Experiences in the Car (pp. 273–294). Springer International Publishing.

49 Wintersberger, P., Schartmüller, C., and Riener, A. (2019). Attentive User Interfaces to Improve Multitasking and Take-Over Performance in Automated Driving: The Auto-Net of Things. International Journal of Mobile Human Computer Interaction, 11(3), 40–58.

50 Yan, F., Eilers, M., Weber, L., and Baumann, M. (2019). Investigating Initial Driver Intention on Overtaking on Rural Roads. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC) (pp. 4354–4359). IEEE.

6.15 Database Indexing and Query Processing

Organizers: Renata Borovica-Gajic, Goetz Graefe, Allison Lee, Caetano Sauer, and Pinar Tözün

Seminar No. 22111

Date: March 13–18, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.3.82

Participants: Carsten Binnig, Thomas Bodner, Peter A. Boncz, Yannis Chronis, Jan Finis, Campbell Fraser, Mhd Yamen Haddad, Stefan Halfpap, Thomas Heinis, David Justen, Andrew Lamb, Allison Lee, Sangjin Lee, Viktor Leis, Alberto Lerner, Thomas Neumann, Anisoara Nica, Danica Porobic, Daniel Ritter, Bart Samwel, Caetano Sauer, Knut Stolze, Pinar Tözün, Nga Tran, Lukas Vogel, Tianzheng Wang, Marcin Zukowski

Remote Participants: Renata Borovica-Gajic, Goetz Graefe

The Dagstuhl Seminar 22111 on "Database indexing and query processing" assembled researchers from industry and academia for the fourth time to discuss robustness issues in database query performance. The seminar gathered researchers around the world working on indexing, storage, plan generation and plan execution in database query processing, and in cloud-based massively parallel systems with the purpose to address the open research challenges with respect to the robustness of database management systems. Delivering robust query performance is well known to be a difficult problem for database management systems. All experienced DBAs and database users are familiar with sudden disruptions in data centers due to poor performance of queries that have performed perfectly well in the past. The goal of the seminar was to discuss the current state-of-the-art, to identify specific research opportunities in order to improve the state-of-affairs in query processing, and to develop new approaches or even solutions for these opportunities, building upon successes of the past Dagstuhl Seminars [1–3]. The organizers (Renata Borovica-Gajic, Goetz Graefe, Allison Lee, Caetano Sauer, and Pinar Tözün) this time attempted to have a focused subset of topics that the participants discussed and analyzed in more depth. From the proposed topics on algorithm choices, join sequences, learned and lightweight indexes, database utilities, modern storage hardware, and benchmarking for robust query processing, the participants formed three work groups: i) one discussing indexing for data warehousing, ii) one discussing robust query operators, and iii) one discussing robust storage architectures. Upon choosing the topics of interest, the organizers then guided the participants to approach the topic through a set of steps: by first considering related work in the area; then introducing metrics and tests that will be used for testing the validity and robustness of the solution; after metrics, the focus was on proposing specific mechanisms for the proposed approaches; and finally the last step focused on the implementation policies. At the end of the week, each group presented their progress with the hope to continue their work towards a research publication. The reports of work groups are presented in the full report.

References

1 Renata Borovica-Gajic, Stratos Idreos, Anastasia Ailamaki, Marcin Zukowski, and Campbell Fraser. Smooth scan: Statistics-oblivious access paths. In Johannes Gehrke, Wolfgang Lehner, Kyuseok Shim, Sang Kyun Cha, and Guy M. Lohman, editors, *ICDE*, pages 315–326. IEEE Computer Society, 2015.

2 Renata Borovica-Gajic, Stratos Idreos, Anastasia Ailamaki, Marcin Zukowski, and Campbell Fraser. Smooth scan: robust access path selection without cardinality estimation. *VLDB J.*, 27(4):521–545, 2018.

3 Martin L. Kersten, Alfons Kemper, Volker Markl, Anisoara Nica, Meikel Poess, and Kai-Uwe Sattler. Tractor pulling on data warehouses. In Goetz Graefe and Kenneth Salem, editors, *DBTest*, page 7. ACM, 2011.

6.16 3D Morphable Models and Beyond

Organizers: Bernhard Egger, William Smith, Christian Theobalt, and Stefanie Wuhrer

Seminar No. 22121

Date: March 20–25, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.3.97

Participants: Thabo Beeler, Volker Blanz, Timo Bolkart, Dan Casas, Bernhard Egger, Victoria Fernandez Abrevaya, James Gardner, Oshri Halimi, Patrik Huber, Marilyn Keller, Ron Kimmel, Adam Kortylewski, Chunlu Li, Ben Mildenhall, Nick Pears, Sami Romdhani, Vincent Sitzmann, Ayush Tewari, Christian Theobalt, Roy Velich, Thomas Vetter, Vanessa Wirth, Stefanie Wuhrer, Tarun Yenamandra

Remote Participants: Michael J. Black, Federica Bogo, Duygu Ceylan, Andrew Fitzgibbon, Tal Hassner, Hao Li, Xiaoming Liu, Arianna Rampini, Emanuele Rodolà, Shunsuke Saito, William Smith, Justus Thies, Maximilian Weiherer, Jiajun Wu, Stefanos Zafeiriou

A total of 63 people were invited to the seminar in the first round of invitations. 39 people attended, with 15 of those attending the seminar virtually. Participants came from both academia and industry and at varying stages of their careers. As this seminar took place at the trailing end of the Covid-19 pandemic, it ran in a hybrid format, and for many attendees, Dagstuhl was the first in-person seminar in several years. Due to the fantastic facilities of the Dagstuhl campus, the hybrid format was a great success, enabling accessible and inclusive communication with remote participants. Daily Covid testing for those in-person ensured that everyone remained safe throughout the week. Eighteen presented their work in around 15–30 minute presentations; an abstract for each talk is included in the full report.

Alongside traditional presentations, many sessions were left available for activities suggested by the seminar participants. These could involve workshops, discussions, presentations, or any other suggested format. During the week, participants could propose plans for the flexible sessions and the structure of the seminar became fixed as activities and topics for the sessions were provided. Summaries for the results of these flexible sessions are contained in the full report. One slot was reserved for a joint group discussion on the ethical concerns of the research we are developing. This resulted in a vivid discussion on the steps we as a community should be taking to encourage the ethical use of the technology we are developing. One of the discussions that received broad support was the design of a cheap, open-source method for collecting camera calibrated illumination environments. This resulted in a Slack channel for the group of interested researchers and the pursuit of an early prototype design. We started the seminar with a short introduction from all participants. Everyone was given one slide to introduce themselves and asked to prepare a question, challenge or goal to discuss during the seminar.

6.17 Framing in Communication: From Theories to Computation

Organizers: Katarzyna Budzynska, Chris Reed, Manfred Stede, and Benno Stein

Seminar No. 22131

Date: March 27–April 1, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.3.117

Participants: Khalid Al-Khatib, Liesbeth Allein, Alexander Bondarenko, Katarzyna Budzynska, Philipp Cimiano, Lucie Flek, Anette Frank, Iryna Gurevych, Ivan Habernal, Annette Hautli-Janisz, Zlata Kikteva, Konrad Kiljan, Christopher Klamm, Marcin Koszowy, Maud Oostindie, Steve Oswald, John Parkinson, Martin Potthast, Aniket Pramanick, Chris Reed, Andrea Rocci, Arno Simons, Wassiliki Siskou, Joanna Skolimowska, Noam Slonim, Manfred Stede, Benno Stein, Henning Wachsmuth

Language is used for many purposes, both private and public. When speech or text is directed to wide audiences, it often aims at influencing stances, opinions, and dispositions of readers. This can be done by relatively transparent, rational argumentation, but also in considerably more subtle ways, by phrasing utterances in such a way that the underlying intent is noticed by readers more in passing – or not consciously at all. This is the realm of "framing", which concerns the careful selecting of the aspects of an event to be reported (those that fit the goal of letting a positive or negative evaluation shine through); the choice of terms that carry an inherent evaluation (e.g., "the frugal four" versus "the stingy four" in recent EU negotiations); and employing stylistic devices that correspondingly support the purpose (e.g., a monotonic versus a lively rhythm). Framing has been studied for quite some time, from many different perspectives, and it has also been covered by popular science books. Under these circumstances, it is not surprising that definitions and emphasis differ quite a bit between and even within disciplines – the notion of framing can itself be framed, too.

The computational research on language processing has addressed some of the linguistic purposes mentioned above: Sentiment analysis and opinion mining are well-established fields; argumentation mining has more recently caught much attention and is in the process of "settling down". Framing, being less transparent at the linguistic surface, has seen only very few attempts at formal modelling so far. The organizers of this seminar are convinced, however, that a computational treatment of framing is a central next step – extending opinion and argument analysis – and its operationalization calls for a deeper understanding of the term and the underlying mechanisms. Before computational theories can be formulated and applications be built, the potential contributions by the various relevant disciplines (sociology, political science, psychology, communication science, and others) should be studied carefully and assessed for potential common ground. This was the first purpose of the seminar, and the second was the follow-up step of developing a roadmap for productive computational research toward the automatic identification of framing in text and speech, and modelling the connection to the underlying reasoning processes.

To accomplish this, the seminar addressed a range of topics, including:

- Argumentation theory, discourse analysis, rhetoric
- Journalism, political science, communication science
- Sociolinguistics, psycholinguistics
- Computational pragmatics and discourse modelling
- Computational social science and social media
- Computational models of argument and debating technologies

6

6.18 Graph Embeddings: Theory meets Practice

Organizers: Martin Grohe, Stephan Günnemann, Stefanie Jegelka, and Christopher Morris

Seminar No. 22132

Date: March 27–30, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.3.141

Participants: Francesco Di Giovanni, Federico Errica, Fabrizio Frasca, Floris Geerts, Martin Grohe, Stephan Günnemann, Nils Kriege, Gitta Kutyniok, Haggai Maron, Christopher Morris, Gaurav Rattan, Bruno Ribeiro, Bastian Rieck, Pascal Schweitzer, Jan Tönshoff

Remote Participants: Pablo Barcelo, Aleksandar Bojchevski, Joan Bruna Estrach, Tina Eliassi-Rad, Matthias Fey, Barbara Hammer, Stefanie Jegelka, Elias Khalil, Benny Kimelfeld, Yaron Lipman, Andreas Loukas, Bryan Perozzi, Siamak Ravanbakhsh, Joshua Robinson, Yizhou Sun, Petar Velickovic, Ulrike von Luxburg, Marinka Zitnik

Graph-structured data is ubiquitous across application domains ranging from chemo- and bioinformatics to image and social network analysis. To develop successful machine learning algorithms or apply standard data analysis tools in these domains, we need techniques that map the rich information inherent in the graph structure to a vectorial representation in a meaningful way – so-called graph embeddings. Designing such embeddings comes with unique challenges. The embedding has to account for the complex structure of (real-world) networks and additional high-dimensional continuous vectors attached to nodes and edges in a (permutation) invariant way while being scalable to massive graphs or sets of graphs. Moreover, when used in supervised machine learning, the model trained with such embeddings must generalize well to new or previously unseen (graph) instances. Hence, more abstractly, designing graph embeddings results in a trade-off between expressivity, scalability, and generalization.

Starting from the 1960s in chemoinformatics, different research communities have worked in the area under various guises, often leading to recurring ideas. Moreover, triggered by the resurgence of (deep) neural networks, there is an ongoing trend in the machine learning community to design invariant/equivariant neural architectures that are capable of dealing with graph- and relational input, both (semi-)supervised and unsupervised, often denoted as graph neural networks. Although successful in practical settings, most of these developments are driven by intuition and empiricism and are geared towards specific application areas. There is no clear understanding of these approaches' limitations and their trade-offs in complexity, expressivity, and generalization. Researchers recently started to leverage connections to graph theory, group theory, logic, combinatorial algorithms, and (algorithmic) learning theory, leading to new theoretical insights and triggering new research in applications. Hence, in this seminar, we aimed to bring together leading applied and theoretical researchers in graph embeddings and adjacent areas, such as graph isomorphism, bio- and chemoinformatics, graph theory, to facilitate an increased exchange of ideas between these communities. Concretely, we aimed to understand what hinders recent theoretical developments being applied in application areas and worked towards a more practical theory. Further, we aimed at understanding the overarching challenges across applications and challenges inherent to specific areas to stimulate directions for further practical and theoretical research.

The seminar bought together 33 researchers from (applied) mathematics, specifically harmonic analysis and (algebraic) topology, (theoretical) computer science, machine learning, bioinformatics, and network science. Eighteen researchers attended remotely owing to the global COVID-19 pandemic. In total, the participants presented 18 talks on their recent progress in a better understanding of graph embeddings, focusing on supervised machine learning, particularly graph neural networks. Many talks dealt with leveraging tools from graph isomorphism testing and related areas such as finite model theory and group theory. In particular, the Weisfeiler-Leman algorithm, a popular heuristic for the graph isomorphism problem, was used to measure the expressivity of the presented algorithms and neural architectures. The consensus was that the above algorithm leads to a too coarse-

grained measure of expressivity, and new notions of expressivity are needed to develop a thorough understanding. Surprisingly, only a few talks dealt with developing a better understanding of generalization, indicating that the research community still lacks an understanding. Notably, Gitta Kutyniok showed how to leverage random graph models and graphons to analyze the generalization error of graph neural networks, while Bruno Ribeiro talked about the connection between causality and out--of-distribution generalization. Further, some talks used methods from (algebraic) topology and their connection to graph theory to devise provably expressive architectures and to better understand common problems with graph neural networks, e.g., the problem of "over-smoothing" of node representations faced when considering deep architectures. Moreover, two talks covered the challenges of applying graph neural networks to biomedical data and industrial applications at Google, respectively, indicating a gap between theoretical results and practical architectures.

Concluding Remarks The seminar was well received, as witnessed by several positive comments from on-site participants. In general, there was an exciting atmosphere at the seminar, particularly among the large number of junior researchers attending the seminar on-site, also witnessed by many lively discussions during on-site talks. However, this was not always the case during online talks, and the active participation of online participants was relatively low. Finally, the organizers wish to express their gratitude to the Scientific Directors of Schloss Dagstuhl – Leibniz Center for Informatics for their support of the seminar.

6

6.19 Symmetric Cryptography

Organizers: Nils Gregor Leander, Bart Mennink, Maria Naya-Plasencia, and Yu Sasaki

Seminar No. 22141

Date: April 3–8, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.4.1

Participants: Subhadeep Banik, Christof Beierle, Ritam Bhaumik, Xavier Bonnetain, Christina Boura, Clémence Bouvier, Anne Canteaut, Patrick Derbez, Orr Dunkelman, Maria Eichlseder, Patrick Felke, Antonio Florez-Gutierrez, Margot Funk, Aldo Gunsing, Ashwin Jha, Pierre Karpman, Daniël Kuijsters, Virginie Lallemand, Eran Lambooij, Nils Gregor Leander, Bart Mennink, Nicky Mouha, Maria Naya-Plasencia, Patrick Neumann, Clara Pernot, Léo Perrin, Shahram Rasoolzadeh, Christian Rechberger, Yann Rotella, Sondre Rønjom, Yu Sasaki, Markus Schofnegger, André Schrottenloher, Yaobin Shen, Tyge Tiessen, Aleksei Udovenko

IT Security plays an increasingly crucial role in everyday life and business. Virtually all modern security solutions are based on cryptographic primitives. Symmetric cryptography deals with the case where both the sender and the receiver of a message use the same key. Due to their good performance, symmetric cryptosystems are highly relevant not only for academia, but also for industrial activities.

We identified the following areas as some of the most important topics on symmetric cryptography at the moment.

Lessons Learnt from NIST Lightweight Cryptography Project. The US National Institute of Standards and Technology (NIST) acknowledged in 2013 the real-world importance of lightweight cryptography, and announced an initiative for standardization. It is expected that the new lightweight standard will not only be used in the US, but rather worldwide.

New Design Strategies. This area deals with the development of symmetric cryptographic primitives and modes that must operate for specific applications, such as STARKs, SNARKs, fully homomorphic encryption, and multi-party computation. These novel applications lead to a paradigm shift in design criteria that we are just starting to understand, both in terms of possible optimizations as well as security impacts.

Quantum-Safe Symmetric Cryptography. For symmetric cryptography, it is short-sighted to expect that cryptanalysis will not improve with the help of quantum computers in the future. It is of importance to understand both the possibility to quantize existing classical attacks, as well as the possibility to perform new types of cryptanalytic attacks using a quantum computer.

Understanding Security Implications from Ideal and Keyless Primitives. Permutation-based cryptography has gained astounding popularity in the last decade, and security proofs are performed in an ideal permutation model. Partly as a consequence of this, the concrete security analysis of the involved primitives has become more difficult. One challenge is to understand (i) to what extent distinguishers impact the security of cryptographic schemes and (ii) what non-random properties of permutations seem likely to be translated into an attack on the full scheme.

Seminar Program

The seminar program consisted of a few short presentations and in-depth group meetings. Presentations were about the above topics and other relevant areas of symmetric cryptography, including state-of-the-art cryptanalytic techniques and new designs. In the full report one can find the list of abstracts for talks given during the seminar.

The research groups were on various topics in symmetric cryptography, all related to one of the above points in one way or another. On the last day of the seminar, the leaders of each group gave brief summaries of achievements. Some teams continued working on the topic after the seminar and started new research collaborations. Here are the summaries of the five groups:

- Group 1 worked and discussed on various problems of provable security, roughly corresponding to one project per person. For three of the projects, the groups had preliminary discussions, and the next step will be to perform the remaining research and investigate the details offline. For two problems, namely improved unforgeability of certain MAC constructions and generic analysis of PRF's and MAC's on 2 public permutations, they advanced quite well and the details will be written down soon after the seminar.
- Group 2 worked on several topics that they plan to continue after the seminar. One was to find good algorithms for detecting the optimal trees of some Boolean functions in the

context of improved key-recovery attacks, and figuring out if we actually need trees, or if we could find or use better partitions that do not correspond to a tree and yet improve the complexity. They also worked on building two attacks on the HALFLOOP construction. They solved the problem of finding structures in linear layers and of decomposing them, and they applied this to Streebog. They also continued developing a new cryptanalysis family; differential meet-in-the-middle attacks. They figured out how to correctly combine it with bicliques, and started working on an application on the construction of SKINNY, which should be comparable if not better than the best known attacks.

- Group 3 worked on several topics related to cryptanalysis, that they plan to continue after the seminar. The studied Tweakable Twine, a tweakable variant of Twine proposed in 2019. They looked at impossible differential distinguishers, but unfortunately they were not able to cover more rounds than in previous work. They also looked at the differential propagation of the cipher. They were able to find a distinguisher that would be established with a probability of 2^{-61}, and they rediscovered a 24-round zero correlation attack in Twine. They have also pointed out several observations on TinyJAMBU, including a method to break the P_b permutation (for 384 rounds) if one can observe collisions during the processing phase. They looked at a paper from 2016 on KATAN that searches for extended boomerang distinguishers. They are implementing the attacks to observe the impact of the middle-round dependencies experimentally. Finally, they looked at (free-start) collisions on Romulus-H and tried to find differential characteristics that are suitable to be used in two SKINNY invocations. One idea would be to use the dependencies to have a collision of a higher probability.
- Group 4 has worked on integral distinguishers on big finite fields. After looking at different topics, this group focused in the following problem: can we find integral distinguishers from the knowledge of some properties of the univariate representation of a function $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$? In other words, they wanted to find some coefficients $(\lambda_0, \ldots, \lambda_{2^n-1})$ in $\mathbb{F}_{2^n}$ such that $\forall x, \sum_{i=0}^{2^n-1} \lambda_i F(\alpha^i X) = 0$. In the particular case where all $\lambda_i \in \mathbb{F}_{2^n}$, this corresponds to finding sets of inputs such that the corresponding outputs sum to zero. They proved that $\sum_{i=0}^{2^n-1} \lambda_i F(\alpha^i X)$ does not contain any term of degree ℓ if and only if $A_\ell = 0$ or $P(\alpha^\ell) = 0$ where $F(X) = \sum_{i=0}^{2^n-1} A_i X^i = 0$. Therefore, they aimed at finding polynomials P which vanish on all α^ℓ when i varies in a given set, and which have the smallest possible number of terms. Indeed, the number of terms of P is the data complexity of the distinguisher. When the only information we have on F is that $A_i = 0$ for all i of weight $\geq d$, then the polynomial P with binary coefficients and with the smallest weight corresponds to the usual distinguisher obtained with higher-order differentials, i.e., $rot(P) = 2^d$. However, if we have more information on A_i, then we can obtain distinguishers with lower data complexity than expected.
- Group 5 looked at a few different topics, quite unrelated to each other. One of them was how to sample binary words of fixed weight (say 200) and length (say 40000) efficiently and in "cryptographic constant time". A possible approach is to use format-preserving encryption, but this turns out to be quite slow compared to alternatives. They eventually slightly revisited an existing method that oversamples w' indices uniformly and independently such that at least w of them are unique with high probability, by proposing a possibly novel and simple constant-time algorithm to extract such a subset of w indices uniformly: write a list (v_i, i) of the w' samples; sort with respect to v_i, mark any duplicate by setting i to infinity; sort with respect to i and keep the w first entries.

  Another topic was the study of the exact differential probability of $1/4$ round of Salsa, or rather computing exactly the probability of any $1/4$ round differential. A "promising" approach would be to use finite automata to parameterize the space of solutions to part of a round, and then iteratively propagate this through the successive steps thereof. They have not implemented this, but one could in principle at least partially rely on some existing tools for the first part. Whether the parameterization would be sufficiently compact to also allow an efficient propagation is not clear yet.

6.20 Recent Advancements in Tractable Probabilistic Inference

Organizers: Priyank Jaini, Kristian Kersting, Antonio Vergari, and Max Welling

Seminar No. 22161

Date: April 18–22, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.4.13

Participants: Alessandro Antonucci, Michael Chertkov, YooJung Choi, Alvaro Correia, Priyank Jaini, Kristian Kersting, Stefan Mengel, Eric Nalisnick, Sriraam Natarajan, Mathias Niepert, Robert Peharz, Xiaoting Shao, Guy Van den Broeck, Antonio Vergari, Andrew G. Wilson

Remote Participants: Marcus A. Brubaker, Cassio de Campos, Nicola Di Mauro, Laurent Dinh, Danilo Jimenez Rezende, Mikko Koivisto, Sara Magliacane, Lilith Francesca Mattei, Denis D. Mauá, Karthika Mohan, David Montalvan Hernandez, Deepak Pathak, Tahrima Rahman, Jakub Tomczak, Aki Vehtari, Max Welling, Yaoliang Yu, Han Zhao

ML models and systems to enable and support decision making in real-world scenarios need to robustly and effectively *reason in the presence of uncertainty* over the configurations of the world that can be observed. *Probabilistic inference* provides a principled framework to carry on this reasoning process, and enables probabilistic *modeling*: a collection of principles to design and learn from data models that are capable of dealing with uncertainty. The main purpose for these models, once learned or built, is to answer *queries* – posed by humans or other autonomous systems – concerning some aspects of the represented world and quantifying some form of uncertainty over it. That is, that is computing some quantity of interest of the probability distribution that generated the observed data. For instance, the mean or the modes of such a distribution, the marginal or conditional probabilities of events, expected utilities of our policies, or decoding most likely assignments to variables (also known as MAP inference, cf. the Viterbi algorithm). Answering these queries reliably and efficiently is more important than ever: we need ML models and systems to perform inference based on well-calibrated uncertainty estimates throughout all reasoning steps, especially when informing and supporting humans in decision making processes in the real world.

For instance, consider a ML system learned from clinical data to support physicians and policy makers. Such a system would need to support *arbitrary queries* posed by physicians, that is, questions that are not known a priori. Moreover, these queries might involve complex probabilistic reasoning over possible states of the world, for instance involving maximization of some probabilities and the ability to marginalize over unseen or not available (missing) attributes like "At what age is a patient with this X-ray but no previous health record most likely to show *any* symptom of COVID-19?", or counting and comparing sub-populations "What is the probability of there being more cases with fever given a BMI of 25 in this county than in the neighboring one?". At the same time, it should guarantee that the uncertainty in its answers, modeled as probabilities, should be faithful to the real-world distribution as uncalibrated estimates might greatly mislead the decision maker.

Recent successes in machine learning (ML) and particularly deep learning have delivered very expressive probabilistic models and learning algorithms. These have proven to be able to induce exceedingly richer models from larger datasets but, unfortunately, at an incredible cost: these models are vastly *intractable* for all but the most trivial of probabilistic reasoning tasks, and they have been demonstrated to provide unreliable uncertainty estimations. In summary, their applicability to real-world scenarios, like the one just described, is very limited.

Nevertheless all these required "ingredients" are within the grasp of several models which we group together under the umbrella name of *tractable probabilistic models*, the core interest of this seminar. *Tractability* here guarantees answering queries *efficiently* and *exactly*. Tractable probabilistic models (TPMs) have a long history rooted in several research fields such as classical probabilistic graphical models (low-treewidth and latent variable models), automated reasoning via knowledge compilation (logical and arithmetic circuits) and statistics (mixture

models, Kalman filters). While these classical TPMs are known to be limited in expressiveness, several recent advancements in deep tractable models (sum-product networks, probabilistic sentential decision diagrams, normalizing flows and neural autoregressive models) are inverting the trend and promising tractable probabilistic inference with little or no compromise when compared to the deep generative models discussed above. It becames then more and more important to have a seminar on these recent successes of TPMs bringing together the different communities of tractable probabilistic modeling at the same table to propel collaborations by defining the goals and the agenda for future research.

These are the major topics around which the seminar brought up the aforementioned discussion:

- Advanced probabilistic query classes
- Deep tractable probabilistic modeling
- Robust and verifiable probabilistic inference
- Exploiting symmetries for probabilistic modelling and applications in science.

Advanced probabilistic query classes

Probabilistic inference can be reduced as computing probabilistic queries, i.e., functions whose output are certain properties of a probability distribution (e.g., its mass, density, mean, mode, etc.) as encoded by a probabilistic model. Probabilistic queries can be grouped into classes when they compute the same distributional properties and hence share the same computational effort to be answered. Among the most commonly used query classes there are complete evidence (EVI), marginals (MAR), conditionals (CON) and maximum a posteriori (MAP) inference. While these classes have been extensively investigated in theory and practice, they constitute a small portion of the probabilistic inference that might be required to support complex decision making in the real-world.

In fact, one might want to compute the probabilities of logical and arithmetic constraints, of structured objects such as rankings, comparing the likelihood and counts of groups of events or computing the expected predictions of discriminative model such as a classifier or regression w.r.t. some feature distribution. Tracing the exact boundaries of tractable probabilistic inference for these advanced probabilistic query classes and devising probabilistic models delivering efficient and reliable inference for them is an open challenge.

Deep tractable probabilistic modeling

A probabilistic model falls under the umbrella name of tractable probabilistic models (TPMs) if it guarantees exact and polytime inference for certain query classes. As different model classes can be tractable representations for different query classes, a *spectrum* of tractable inference emerges. Typically, this create a tension with the extent of a model class supporting a larger set of tractable query classes, and its *expressive efficiency*, i.e., the set of functions it can represent compactly.

Recent deep generative models such as generative adversarial networks (GANs), regularized and variational autoencoders (VAEs) fall out of the TPM umbrella because they either have no explicit likelihood model or computing even the simplest class of queries, EVI, is hard in general. In fact, despite their successes, their inference capabilities are severely limited and one has to recur to approximations. However, the approximate inference routines available so far (such as the evidence lower bound and its variants) do not provide sufficiently strong guarantees on the quality of the approximation delivered to be safely deployed in real-world scenarios.

On the other hand, classical TPMs from the probabilistic graphical model community support larger classes of tractable queries comprising MAR, CON and MAP (to different extents based on the model class). Among these there are: i) low or bounded-treewidth probabilistic graphical models that exchange expressiveness for efficiency; ii) determinantal point processes which allow tractable inference for distributions over sets; iii) graphical models with high girth or weak potentials, that provide bounds on the performance of approximate inference methods; and iv) exchangeable probabilistic models that exploit symmetries to reduce inference complexity.

A different prospective on tractability is brought by models compiling inference routines into efficient computational graphs such as arithmetic circuits, sum-product networks, cutset networks and probabilistic sentential decision diagrams have advanced the state-of-the-art inference performance by exploiting context-specific independence, determinism or by exploiting latent variables. These TPMs, as well as many classical tractable PGMs as listed above, can be cast under a unifying framework of probabilistic circuits (PCs), abstracting from the different graphical formalisms of each model. PCs with certain structural properties support tractable MAR, CON, MAP as well as some of the advanced query classes touched in the previous topic item. Guy Van den Broeck gave a long talk on the first day of the seminar to set the stage for participants for viewing tractable probabilistic models from the lens of probabilistic circuits.

More recently, the field of neural density estimators has gained momentum in the tractable probabilistic modeling community. This is due to model classes such as normalizing flows and autoregressive models. Autoregressive models and flows retain the expressiveness of GANs and VAEs, by levering powerful neural representations for probability factors or invertible transformations, while overcoming their limitations and delivering tractable EVI queries. As such, they position themselves in the spectrum of tractability in an antithetic position w.r.t. PCs: while the latter support more tractable query classes, the former are generally more expressive. On the first day of the seminar, Marcus Brubaker introduced these models to the seminar participants in a long talk. It is an interesting open challenge to combine TPM models from different regions of such a spectrum to leverage the "best of different worlds", i.e., increase a model class expressive efficiency while retaining the largest set of supported tractable query classes as possible. The first day subsequently ended with a lively open discussion on the differences between TPMs and Neural Generative Models and what advantages and lessons they can provide the other models.

Robust and verifiable probabilistic inference

Along exactness and efficiency, one generally requires inference routines to be robust to adversarial conditions (noise, malicious attacks, etc.) and to be allow exactness and efficiency to be formally provable. This is crucial to deploy reliable probabilistic models in real-world scenarios (cf. other topic). Recent advancements in learning tractable and intractable probabilistic models from data have raised the question if the learned models are just exploiting spurious correlations in input space, thus ultimately delivering an unfaithful image of the probability distribution they try to encode. This raises several issues, as in tasks like anomaly detection and model comparison, which rely on correctly calibrated probabilities, one can be highly

mislead by such unfaithful probabilistic models. Furthermore, one might want to verify a priori or ex-post (e.g., in presence of adversarial interventions) if one probabilistic inference algorithm truly guarantees exact inference. Questions like this have just very recently been tackled in a formal verification setting, where proofs of the correctness of inference can be verified with less resources than it takes to execute inference.

Over the course of the seminar, through informal discussions and formal talks by the participants discussed the above mentioned issues in tractable probabilistic inference through topics such as Bayesian Deep Learning, Incorporating symmetries in probabilistic modelling using equivariance with applications in sciences, explainable AI etc.

Overall, the seminar produced numerous insights into how efficient, expressive, flexible, and robust tractable probabilistic models can be built. Specially, the discussions and talks at the seminar spurred a renewed interest in the community to:

- develop techniques and approaches that bring together key ideas from several different fields that include deep generative models, probabilistic circuits, knowledge compilation, and approximate inference.
- create bridges between researchers in these different fields and identify ways in which enhanced interaction between the communities can continue.
- generate a set of goals, research directions, and challenges for researchers in these field to develop robust and principled probabilistic models.
- provide a unified view of the current undertakings in these different fields towards probabilistic modelling and identifying ways to incorporate ideas from several fields together.
- develop a new systematic and unified set of development tools encompassing these different areas of probabilistic modelling.

6.21 Urban Mobility Analytics

Organizers: David Jonietz, Monika Sester, Kathleen Stewart, Martin Tomko, and Stephan Winter

Seminar No. 22162

Date: April 18–22, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.4.26

Participants: Vanessa Brum-Bastos, Hao Cheng, Tao Cheng, David Doerr, Christian Eichenberger, Cheng Fu, Martin Lauer, Dirk Christian Mattfeld, Alexandra Millonig, Edoardo Neerhut, Moritz Neun, Daniel Nüst, Erik Nygren, Maya Sekeran, Monika Sester, Martin Tomko, Stephan Winter, Yanan Xin

Remote Participants: Andris Clio, Ioannis Giannopoulos, Anita Graser, Ivan Majic, Michael Nolting, Luca Pappalardo, Chiara Renso, Piyushimita Vonu Thakuriah

Seminar 22162 addressed recent trends in urban mobility analytics that are shaping the information available to transport planners, operators, and travellers. Seminar participants were particularly discussing how information can be provided that supports the critical transformation of urban mobility towards climate neutrality and other sustainability goals, i.e, that supports to change mobility behaviour.

The trends identified for this seminar were, on one hand, the rise of deep learning methods for massive data analytics, and on the other hand the emerging digital divide between those having massive data and those who haven't, which, in short, forms the challenges of academia for reproducible research. Massive data on urban mobility is collected by industry and transport authorities, with limited access outside, for various reasons. Also, the research and development capacity behind the closed doors of large transnational companies – especially in the platform economy – is arguably faster than the typical PhD process.

These challenges and opportunities were shaping the discussions where participants split into working groups on (a) *ethics and the social good* – how can information trigger change in mobility behaviour; (b) *methods and explainability*; (c) *benchmarking and datasets*; and (d) *applications*.

The seminar had quite a diversity of participants, which was inspiring in all the discussions. Participants from industry gave talks about what happened behind their 'closed doors', and further tutorials were introducing datasets, the principle of reproducible research, and European funding opportunities.

The industry partners showed great interest in collaboration with academia, however, the problem of data sharing was still considered as paramount. There are trends to open certain kinds of data, e.g., in aggregated form, or simulated data, or only based on contracts with certain institutions. Still, open data sharing remains to be a challenge.

6.22 Digital Twins for Cyber-Physical Systems Security

Organizers: Mohammad Al Faruque, Alvaro Cárdenas Mora, Simin Nadjm-Tehrani, and Edgar Weippl

Seminar No. 22171

Date: April 24–29, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.4.54

Participants: Ali Abbasi, David Allison, Magnus Almgren, Alvaro Cárdenas Mora, Katharina Ceesay-Seitz, Matthias Eckhart, Andreas Ekelhart, Helge Janicke, Marina Krotofil, Martina Maggio, Simin Nadjm-Tehrani, Miroslav Pajic, Awais Rashid, Francesco Regazzoni, Anne Remke, Henrik Sandberg, Anne-Kathrin Schmuck, Nils Ole Tippenhauer, Edgar Weippl, Mark Yampolskiy

In the light of the increasing digitization and move toward Industry 4.0 [1], cyber security becomes more and more important for cyber-physical systems (CPSs). The advanced computation, communication, and control capabilities of CPSs lead to a wider attack surface and greater exposure to security flaws. Furthermore, the added complexity puts a considerable burden on security professionals, who have to ensure that the CPSs are adequately protected against adversaries throughout the entire lifecycle. As a matter of fact, designing holistic security measures is a significant ongoing challenge for academia and industry alike. Thorough security testing during the engineering- and, particularly, the operation phase is often not feasible. The development of custom CPS testbeds is complicated, expensive, and time-consuming due to high hardware costs, space constraints, and complex dependencies between components [2]. Past attempts to conduct penetration tests directly on live systems led to unintended system behavior, putting human workers in significant danger and causing a disruption of production lines [3]. In addition to regular security testing, adequate countermeasures need to be implemented in response to newly discovered vulnerabilities that emerge during operation or if the CPS is already under attack. However, the steadily increasing sophistication of cyberattacks calls for more effective intrusion detection and prevention techniques. On top of that, new mechanisms to test and evaluate attack response strategies in a controlled setting are required.

A digital twin, that is, a virtual replica of a real system, was originally envisioned for similar, yet non-security-related purposes: The life of a spacecraft is virtually mirrored through high-fidelity simulations and sensor updates to detect anomalies and safely test mitigation options such that degradation can be reduced and damages prevented [4]. This idea was picked up by the industrial informatics community, whose members implemented the digital-twin concept in various CPS applications for monitoring, lifecycle management, and decision support [5–7]. In the past few years, researchers have also shown interest in utilizing digital twins for security-enhancing purposes [8–13]. Although the definition of what constitutes a digital twin in the context of cybersecurity differs in the literature, its main application areas seem to be clear: Virtually replicated systems by means of emulation, simulation, and modeling technologies, coupled with real-time or historical data flows, might be used to improve security testing, intrusion detection, and attack recovery. However, fundamental research questions and challenges remain before digital twins can be applied for security-enhancing purposes. Furthermore, concerns have been raised about the potential security threats associated with the digital-twin concept [14].

Thus, the primary goal of this Dagstuhl Seminar was to lay the foundation for future interdisciplinary collaboration on digital-twin research for CPS security. The interdisciplinary character of this novel research area is reflected in its origin. As already indicated, the notion of using "twins" originally emerged from the space industry [6], gained wider adoption by the industrial informatics community [5–7], and was eventually applied with the objective of attaining security improvements [8–13]. For this reason, the seminar has brought together 20 researchers with backgrounds in computer security, control theory, automation engineering, and data science. Inspired by the concept's promised security improvement potential, the seminar was structured along three different themes:

Foundations of Security-focused Digital Twins. This theme was motivated by the lack of clarity around the digital-twin concept. Therefore, the purpose of this theme was to develop a common understanding of what a digital twin in the context of security is, how it can be defined, and how it relates to existing concepts, such as cyber ranges, data-driven models, and honeypots. Closely tied to this theme were discussions on methods for digital-twin implementation, including

(i) emulating systems and simulating physical processes,
(ii) knowledge retrieval for digital-twin generation in greenfield and brownfield environments, and
(iii) synchronizing digital twins with their physical counterparts.

Intrusion Detection. The objective of this theme was to explore intrusion detection as a potential use case for digital twins. Assuming that the digital twin is built from a benign specification such that legitimate behavior is exhibited when executed in sync with its counterpart, any deviations observed on the logic, network, and physics layers could indicate malicious activity. Building on this idea, participants discussed how digital twins can serve as a foundation for such behavior-specification-based intrusion detection systems (IDSs) that possess physics- and process-aware capabilities. Moreover, discussions touched on how digital twins can be used for data generation purposes to improve the training phase of (semi-)supervised learning approaches that are employed in behavior-based IDSs.

Attack Response Mechanisms. The last theme was associated with research questions on implementing proactive and reactive attack response strategies, which may represent another use case of digital twins. Proactive security measures can prevent cyber-physical attacks in the face of imminent threats when new vulnerabilities in the CPS are discovered. On the other hand, reactive responses to an attack can be initiated to control damage by ensuring that the physical system maintains a safe state. In this context, questions were raised about how the digital-twin concept can help in designing attack-resilient CPS architectures and response strategies for control systems. This theme highlighted the benefits and challenges of using digital twins to test countermeasures in a simulated environment and assess their effects.

The program started with a welcome session that provided an opportunity for participants to get to know one another. Furthermore, the organizers used this session to share information about the seminar program and explain key terms to participants who were not au fait with the terminologies used by different communities. Over the five days, 14 participants gave lightning talks that focused on the following topics:

- building blocks for digital-twin construction, including emulating and simulating CPS components, data-driven approaches and semantic technologies, synchronization mechanisms,
- reverse engineering programmable logic controllers, deception technology (e.g., honeypots), security testbeds,
- attack detection in CPSs, featuring physics-based, data-driven, and process-aware techniques,
- attack-resilient control using different tools for risk mitigation (viz., prevention, detection, and treatment),
- various aspects of dataset availability in CPS research (e.g., attack simulation, data collection, evaluation, and validation), and
- digital-twin use cases for the safety-related system development lifecycle.

The lightning talk sessions offered each speaker 15 minutes to present new perspectives and talk about current challenges in CPS security. The highly interdisciplinary setting and stimulating presentations given by participants resulted in active discussions, which were carried on in the breakout sessions.

The afternoons of Monday, Tuesday, and Wednesday were used for breakout sessions to give participants the opportunity to work together on research issues of common interest. Based on the discussions that took place on Monday after the session on bridging the disciplinary gap, we identified the following topics of interest to be explored by working groups:

(i) conceptualization of the digital twin for cyber-physical systems security, and
(ii) attack recovery for control systems.

Participants who worked on the former topic discussed characteristics that digital twins need to have to be useful for security applications, while those who focused on the latter topic investigated strategies in the context of control theory to respond to attacks in a reactive manner.

The seminar received very positive feedback from participants, who also expressed strong interest in future editions. In addition, several invitees, who were forced to cancel their participation at short notice due to the SARS-CoV-2 pandemic, have shown great interest in follow-up events. Thus, we believe that this Dagstuhl Seminar should be repeated in the future. A second edition would be worthwhile to investigate open problems concerning system emulation. These issues could be addressed in a future follow-up seminar if more participation from the embedded systems and systems security communities is achieved.

As the organizers, we would like to thank everyone who attended this seminar for their interesting talks, the thought-provoking questions, and the fruitful contributions that led to a highly collaborative atmosphere for scientific discussions. We also would like to express our sincere gratitude to the scientific and administrative staff of Schloss Dagstuhl for their outstanding support that made this seminar possible.

References

1 Henning Kagermann, Johannes Helbig, Ariane Hellinger, and Wolfgang Wahlster. Recommendations for implementing the strategic initiative INDUSTRIE 4.0 – securing the future of german manufacturing industry. Final report of the Industrie 4.0 working group, acatech – National Academy of Science and Engineering, München, April 2013.

2 Benjamin Green, Anhtuan Lee, Rob Antrobus, Utz Roedig, David Hutchison, and Awais Rashid. Pains, gains and PLCs: Ten lessons from building an industrial control systems testbed for security research. In *10th USENIX Workshop on Cyber Security Experimentation and Test (CSET 17)*, Vancouver, BC, 2017. USENIX Association.

3 David Duggan, Michael Berg, John Dillinger, and Jason Stamp. Penetration testing of industrial control systems. *Sandia National Laboratories*, 2005.

4 Mike Shafto, Mike Conroy, Rich Doyle, Ed Glaessgen, Chris Kemp, Jacqueline LeMoigne, and Lui Wang. Draft modeling, simulation, information technology & processing roadmap. *Technology Area*, 11, 2010.

5 Elisa Negri, Luca Fumagalli, and Marco Macchi. A review of the roles of digital twin in CPS-based production systems. *Procedia Manufacturing*, 11:939 – 948, 2017. 27th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017, 27–30 June 2017, Modena, Italy.

6 Roland Rosen, Georg von Wichert, George Lo, and Kurt D. Bettenhausen. About the importance of autonomy and digital twins for the future of manufacturing. *IFAC-PapersOnLine*, 48(3):567 – 572, 2015. 15th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2015.

7 Werner Kritzinger, Matthias Karner, Georg Traar, Jan Henjes, and Wilfried Sihn. Digital twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, 51(11):1016 – 1022, 2018. 16th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2018

8 Matthias Eckhart and Andreas Ekelhart. *Digital Twins for Cyber-Physical Systems Security: State of the Art and Outlook*, chapter 14, pages 383–412. Springer International Publishing, Cham, 2019.

9 Mariana Segovia and Joaquin Garcia-Alfaro. Design, modeling and implementation of digital twins. *Sensors*, 22(14), 2022.

10 Marietheres Dietz and Gunther Pernul. Unleashing the digital twin's potential for ICS security. *IEEE Security & Privacy*, 18(4):20–27, July 2020.

11 David Holmes, Maria Papathanasaki, Leandros Maglaras, Mohamed Amine Ferrag, Surya Nepal, and Helge Janicke. Digital twins and cyber security – solution or challenge? In *2021 6th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM)*, pages 1–8, September 2021.

12 Rajiv Faleiro, Lei Pan, Shiva Raj Pokhrel, and Robin Doss. Digital twin for cybersecurity: Towards enhancing cyber resilience. In Wei Xiang, Fengling Han, and Tran Khoa Phan, editors, *Broadband Communications, Networks, and Systems*, pages 57–76, Cham, 2022. Springer International Publishing.

13 Abhishek Pokhrel, Vikash Katta, and Ricardo Colomo-Palacios. Digital twin for cybersecurity incident prediction: A multivocal literature review. In *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops*, pages 671–678, New York, NY, USA, 2020. Association for Computing Machinery.

14 Cristina Alcaraz and Javier Lopez. Digital twin: A comprehensive survey of security threats. *IEEE Communications Surveys & Tutorials*, 2022.

6.23 Technologies to Support Critical Thinking in an Age of Misinformation

Organizers: Andreas Dengel, Laurence Devillers, Tilman Dingler, Koichi Kise, and Benjamin Tag

Seminar No. 22172

Date: April 24–27, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.4.72

Participants: Chris Coward, Henriette Cramer, Andreas Dengel, Tilman Dingler, David Eccles, Nabeel Gillani, Koichi Kise, Dimitri Molerov, Albrecht Schmidt, Gautam Kishore Shahi, Benjamin Tag, Roger Taylor, Niels van Berkel, Andrew Vargo, Eva Wolfangel

Remote Participants: Susanne Boll, Nattapat Boonprakong, Laurence Devillers, Stephan Lewandowsky, Philipp Lorenz-Spreen, Emma Spiro

The Dagstuhl Seminar on "Technologies to Support Critical Thinking in an Age of Misinformation" ran over a course of three days in April 2022. Each day focused on one specific aspect of the problem of Misinformation and the role technologies play in its worsening and mitigation.

Day 1 put the overall seminar goals and an introduction to the topic into its focus. All participants introduced themselves and gave a concrete example of an important challenge they have identified. The collected challenges were organized and later used as core challenges for group work activities, here Regulations/Policies, Human Factors and Platforms, and Critical Thinking. Over the course of the three days three groups worked on defining challenge statements (Day 1), ideas to solve the issue (Day 2), and concrete Research Questions and Project/Collaboration proposals (Day 3).

The theoretical underpinnings of all group discussions and activities were provided by a series of presentations that were topically organized. Day 1 was centered around how the problem of misinformation has evolved and why misinformation is so successful these days. A historical overview was given by keynote speaker Prof. Emma Spiro, which concluded with the key insights that Networks and platforms shape information flow and that attention dynamics matter. The second keynote talk of the day was given by Prof. Andreas Dengel that put light on the crucial role that images and their power to convey information that is tainted with emotional information, and how technology (e.g, CNNs) can be used to detect those, classify them, and can potentially correct them.

On day 2, the participants zeroed in on the role technology plays. Session 1, started with a keynote by Prof. Niels van Berkel on the role of Artificial Intelligence, and Human-AI interaction. Looking at Technology, Society, and Policy on a larger scale, van Berkel identified the core issue that there exists a lack of literacy on the tech side as well as on the regulatory side, a potential consequence of the lack of qualified tech personnel on regulatory bodies. Keynote 2, by Prof. Laurence Devillers, looked at how technology is used to misinform, deceive, and change public opinion, while proposing solutions, such as Nudging and Boosting techniques, how Human-Ai interaction should be better understood, and how research and industry must work together to mitigate the problem of lacking literacy. In session 2 of the day, Prof. Albrecht Schmid led an open, provocative discussion that served as a brainstorming session for the upcoming group work, mainly focussing on the role of platforms and technology. The third keynote was given by Prof. Stephen Lewandowsky who gave a detailed account of the role of human cognition and the larger impact of misinformation on democratic societies. He identified pressure points and proposes countermeasures that are effective but need to be scaled up through improved and coordinated cross-country regulation. Day 2 ended with a Misinformation Escape Room group activity (demo), led by Dr. Chris Coward, which aims at teaching players the power of misinformation and the complexity of the problem.

Day 3 featured the keynote by Roger Taylor which strongly focussed on the way misinformation is regulated globally, and how regulatory frameworks (Digital Service Act) and effective regulation can help to mitigate the misinformation problem. As an advisor to the UK government, and an expert in responsible AI programs and data ethics, Roger Taylor put a light on pain points in the bureaucracy and the misaligned aims of technology development and research, and politics.

6.24 Computational Metabolomics: From Spectra to Knowledge

Organizers: Corey Broeckling, Timothy Ebbels, Ewy Mathé, and Nicola Zamboni

Seminar No. 22181

Date: May 1–6, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.5.1

Participants: Sebastian Böcker, Corey Broeckling, Roman Bushuiev, Timothy Ebbels, Soha Hassoun, Carolin Huber, Katerina Kechris, Oliver Kohlbacher, Anneli Kruve, Tytus Mak, Lennart Martens, Ewy Mathé, María Eugenia Monge, Steffen Neumann, Louis-Felix Nothias, Rui Pinto, Tomáš Pluskal, Hannes Röst, Juho Rousu, Francesco Russo, Adriano Rutz, Michael Andrej Stravs, Panteleimon Takis, Justin van der Hooft, Cecilia Wieder, Egon Willighagen, Michael Anton Witting, Mitja Zdouc

Metabolomics is the study of small molecules in living systems, including those which generate the energy to sustain life, those that form the building blocks of macromolecules such as DNA, as well as some originating outside the living system such as pollutants. Biologically, this field is of increasing importance due to its strong connection to organism function. Metabolomics is rapidly expanding with significant advances in both measurement technology (e.g., mass spectrometry, chromatography, NMR spectroscopy) and informatics approaches. The amount and complexity of data routinely exceeds the capacity of typical software and other computational systems used in bioanalytical labs and there is an ongoing and increasingly acute need for improvements in computational, informatics and statistical/machine learning approaches to make sense of it all.

This seminar, the fourth in the series on computational metabolomics, continued some themes previously well developed, and explored many new ones. A good example of the former is the problem of how to use mass spectral data to annotate (putatively identify) the 1000s of unknown metabolites typically observed in routine assays. Another example would be the discussion of new developments in dealing with Data Independent Acquisition which has diversified considerably in the last 5 years. Many new directions were also discussed. For instance, the question of pathway analysis – how to generate semi-automated interpretation of metabolomics data on the level of groups of molecules working together in biological processes – is becoming more prominent as larger annotated datasets become available. Another new direction was "metabolproteomics", looking at the diverse array of interactions between metabolites and proteins, in particular in how metabolite derived post-translational modifications of proteins can be picked up in annotation pipelines. Other discussions focused on software aspects such as visualization of chemical space (a key problem in designing effective software tools) and the generation/curation of high quality data for benchmarking new informatics algorithms. A session on extended metabolic models looked at ways to link data and prediction tools from protein function studies to metabolites in order to gain new knowledge of unknown metabolic pathways. From a data generation technology perspective, while mass spectrometry (MS) dominated as expected (e.g., sessions on MS spectral quality requirements, fragmentation trees etc), the seminar extended beyond previous ones into a discussion of NMR data processing and modeling. Open databases, repositories and knowledge representation also featured their own discussions including Wikidata, CxSMILEs and Wikipathways/RaMP-DB. Finally the important issue of integrating metabolomic data with other relevant data types (e.g., genomics, proteomics etc) was discussed.

The seminar organization followed a similar flexible format to the previous one, where topics were both suggested in advance and brainstormed on the Monday. The whole group participated in brainstorming and prioritization and this was further refined each morning of the meeting. Parallel discussions were organized with the aim to minimize clashes in individual interests and at the end of each morning/afternoon session a plenary feedback session was held to disseminate the main discussion points to the whole group. Evening sessions were very popular and covered a wide range of topics. Overall the seminar was felt to be one of the most successful yet, highlighting the growing importance of computational metabolomics as a field in its own right and emphasizing the need for further meetings to address the important problems in this exciting area of research.

6.25 Estimation-of-Distribution Algorithms: Theory and Applications

Organizers: Josu Ceberio Uribe, Benjamin Doerr, Fernando Lobo, and Carsten Witt

Seminar No. 22182

Date: May 1–6, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.5.17

Participants: Carlos Andreu Vilarroig, Mayowa Ayodele, Alexander Brownlee, Josu Ceberio Uribe, Benjamin Doerr, Nikolaus Hansen, Ekhine Irurozki, Ata Kaban, Martin S. Krejca, Per Kristian Lehre, Johannes Lengler, Manuel López-Ibáñez, John McCall, Vicente Perez Soloviev, Jonathan L. Shapiro, Thomas Stützle, Dirk Sudholt, Andrew M. Sutton, Carsten Witt

The seminar "Estimation-of-Distribution Algorithms: Theory and Practice" on May 2–6, 2022 brought together 19 international experts in estimation-of-distribution algorithms (EDAs). Their research ranged from a theoretical perspective, e.g., runtime analysis on synthetic problems, to an applied perspective, e.g., solutions of industrial optimization problems with EDAs.

The main aim of the seminar was to narrow the gap between theory and practice in EDAs by bringing together researchers from both sides and stimulating interaction. We facilitated this interaction through longer introductory talks, e. g., on theoretical analyses of EDAs, regular conference-style talks, short flash talks presenting open problems and stimulating discussions and, last but not least, through various breakout sessions and group work. After each talk, we scheduled ample time for discussion, and even adapted the schedule when discussions had gained momentum and took longer than expected. On the last day of the seminar, all participants joined a 1-hour plenum discussion summarizing the findings of the seminar, discussing open problems and identifying further research topics.

We believe that the seminar has achieved its main aims by making a step towards narrowing the gap between theory and practice in EDAs. This is witnessed by the high number of spontaneous talks given by the participants, allowing almost everyone to present his/her perspective on EDAs, high and stable attendance of participants at talks and group sessions, and lively discussions after basically every talk as well as in the dedicated discussion fora. Several participants shared the feedback with us that they learned new aspects of EDAs during the seminar and had increased their understanding of what the other side (theory/practice) was interested in.

The seminar also identified several open problems related to the design, analysis, and application of EDAs. Details can be found in the summary of the concluding plenum discussion (see the full report).

Finally, several participants reported to the organizers that the seminar had helped them to understand current challenges in the theory and practice of EDAs, that they had extended their professional network, and drawn inspiration from the seminar for new research ideas. They also expressed interest in attending a similar seminar in the future.

The only downside of this seminar was that several colleagues, most notably from China, could not attend because of the still ongoing sanitary crisis.

6

6.26 Visual Text Analytics

Organizers: Christopher Collins, Antske Fokkens, Andreas Kerren, and Chris Weaver

Seminar No. 22191

Date: May 8–13, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.5.37

Participants: Richard Brath, Angelos Chatzimparmpas, Christopher Collins, José Angel Daza Arévalo, Mennatallah El-Assady, Alex Endert, Jean-Daniel Fekete, Antske Fokkens, Yoav Goldberg, Pantea Haghighatkhah, Daniel A. Keim, Andreas Kerren, Johannes Knittel, Kostiantyn Kucher, Ross Maciejewski, Narges Mahyar, Christofer Meinecke, Shimei Pan, Carita Paradis, Barbara Plank, Vasiliki Simaki, Maria Skeppstedt, Pia Sommerauer, Bettina Speckmann, Hendrik Strobelt, Nicole Sultanum, Tatiana von Landesberger, Chris Weaver

Introduction

Visualizing textual information is a particularly challenging area of information visualization and visual analytics research. The types of data processing and analytic algorithms differ greatly from tabular or geospatial data, and the visualization techniques have additional constraints to consider, including the provision of context for text fragments of similar or different size and structure, depicting embeddings and high dimensional representations, and ensuring legibility of text incorporated into visualizations. The wide variation in the data is accompanied by the difficulties in inferring the semantic meaning of ambiguous terms, or determining the referencing between subsequent statements.

This Dagstuhl Seminar succeeded in bringing together researchers from the visualization, natural language processing (NLP), and machine learning communities, with domain experts from several text-related research areas, to identify the most pressing and promising open problems for collaborative research. This truly interdisciplinary approach offered new opportunities to capitalize on existing knowledge and recent developments across all involved disciplines. Discussions in the seminar were comprehensive, focusing on visual text analytics with the goal to provide an application-oriented research agenda.

The seminar coalesced an international community of experts from different disciplines around a research roadmap for the next 5–10 years, as documented through working group reports. The seminar generated a series of research questions which serve as a call to action to the wider community. The unique and contained setting of Schloss Dagstuhl facilitated new cross-disciplinary collaborations and allowed us to lay the groundwork for productive future collaborations, including a planned special issue of the Information Visualization journal.

Seminar Themes The following high-level themes were discussed during the seminar. The seminar allowed attendees to critically reflect on current research efforts, the state of field, and key research challenges today. Participants also were encouraged to demonstrate their system prototypes and tools relevant to the seminar topics. As a result of the first working groups, as well as impromptu demonstrations and discussions, the actual seminar discussion topics evolved and we established a second set of working groups halfway through the week, cf. Section 6 of the full report.

- **Data Sources and Diversity** What is the current landscape of the application fields and data domains? What are the data gaps? Can existing approaches be generalized?
- **Model Explainability and Interpretability** Can we provide more sophisticated visualizations to study how language models learn or what information they represent?
- **Evaluation and Experimental Designs** Which experimental methods best support the evaluation of techniques and processes for visualizing text information?
- **Interaction Design** What design opportunities are unique to, or more pressing, for text data? How can interaction principles be applied to any underlying NLP as well?
- **Toolkits and Standards** What success stories regarding existing text visualization approaches and systems can we learn from? What is needed?
- **TextVis Literacy** Visual text analytics can be applied across a wide variety of domains. How do we make techniques easy to learn and to interpret correctly?

Outcomes

The Dagstuhl team performed an evaluation at the end of the seminar week. The results of this survey (scientific quality, inspiration to new ideas/projects/research/papers, insights from

neighboring fields, ...) were universally very good to excellent. Only a few single improvements were proposed by participants, for example, having longer breaks and mixing up the demo presentations with the other parts of the schedule. Another suggestion was to skip the intermediate group report session because it interrupted the group work.

At the end of the week the organizers agreed to proceed to arrange for a special issue of the journal Information Visualization, which will have an open call but with the intent to include any extended works resulting from the seminar. In addition, several working groups with more "position paper" style reports planed to submit these to well-read venues accepting of editorial works which motivate the research community.

Remaining Challenges in Visual Text Analytics
Not all topics identified during the seminar could be addressed in the working groups and might be left for a future Dagstuhl Seminar on a similar subject area. In the following, we briefly list those topics and open problems (more are surely existing that are not mentioned here):

- *Interaction Design:* Interaction methodologies as part of any visual text analytics approach were in the focus of several working groups. A more systematic classification and evaluation of interaction techniques that are unique for text data would be useful for future developments.
- *Toolkits and Standards:* Even if many toolkits and existing standards were discussed in the seminar, a proper and comprehensive analysis of those is still missing that would be beneficial for users and developers of visual text analytics systems.
- *TextVis Literacy:* This topic is important to broaden the use of visual text analytics techniques in general and should be studied deeper in the future.
- *Focus on Text Data Aspects:* The consideration of data diversity, data fusion, and data organization in context of visual text analytics might be an interesting topic for further discussion.
- *Focus on Specific NLP and ML Methods:* The increasing number of specific/novel analytical methods (such as transfer learning or others) raise the need for specific answers from the visual text analytics community.

Acknowledgments

We would like to thank all participants of the seminar for the lively discussions and contributions during the seminar as well as the scientific directorate of Dagstuhl Castle for giving us the possibility of organizing this event. Angelos Chatzimparmpas gathered the abstracts for the overview of the invited talks, the tool demos, and the working groups in the full report, in Section 4, Section 5, and Section 6, respectively. Once more, we are thankful to all the attendees for agreeing to compose the abstract texts and timely provide them to us in order to write this executive summary. Last but not least, the seminar would not have been possible without the great help of the staff at Dagstuhl Castle. We acknowledge all of them and their assistance.

6

6.27 Dynamic Traffic Models in Transportation Science

Organizers: Martin Gairing, Carolina Osorio, Britta Peis, and David Watling

Seminar No. 22192

Date: May 8–13, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.5.92

Participants: Joseph Chow, Richard Connors, Katharina Eickhoff, Miriam Fischer, Gunnar Flötteröd, Martin Gairing, Yiannis Giannakopoulos, Lukas Graf, Tobias Harks, Martin Hoefer, Takamasa Iryo, Max Klimm, Ekkehard Köhler, Jannik Matuschke, Haruko Nakao, Neil Olver, Carolina Osorio, Dario Paccagnan, Britta Peis, Rahul Savani, Guido Schäfer, Daniel Schmand, Alexander Skopalik, Martin Skutella, Martin Strehler, Bernhard von Stengel, David Watling

Traffic assignment models are crucial for transport planners to be able to predict the traffic distributions – and thereby the congestion, environmental and social impacts – of transport policies, for example in the light of possible changes to the infrastructure (e.g., traffic light controls), to the transport services offered (e.g., the provision of public transport fleets and frequencies), or to the prices imposed on travellers (e.g., road tolls, public transport fares). The prevailing approaches used in the transportation science literature to predict such distributions can be roughly classified into mathematical approaches based on *dynamic traffic assignment models* (DTA) (using the methodology of flows over time) and *simulation-based approaches* (using large-scale microsimulations). The striking advantage of microscopic simulations over DTA models is that the latter usually ignore practically relevant side-constraints such as horizontal queueing of vehicles, the feedback of changing network conditions on user behavior, flexible departure time choice, mode choice, activity schedule choice, and such. Current simulation tools integrate all these dimensions and many more. The increase in model complexity inherent in simulations, however, is not matched by the existing DTA theory. For most simulation software, there is no "mathematical proof of concept" such as formal proofs for the following fundamental properties:

- a (dynamic) equilibrium always exists,
- an equilibrium is unique,
- an equilibrium is efficiently computable,
- there is a smooth function of equilibria with respect to parameters (such as street capacities, tolls, etc.)

In this seminar we brought together, for the third time, leading researchers from three different communities (Simulations, Dynamic Traffic Assignment and Algorithmic Game Theory) in order to bridge the gap between complex simulation based models and the existing theory. In doing so, we build on the progress made in the first two seminars, and at the same time addressed the emerging real-life challenge of modeling the new kinds of 'mobility service' that are increasingly envisaged for the future, such as intelligent/dynamic ride-sharing and car-sharing, through to fully autonomous vehicles, provided potentially by a variety of competing operators.

The purpose of this third seminar was to build on the momentum, collaborations and new directions identified at the previous two seminars, following on from the impacts identified earlier. Similarly to the first two seminars, one objective was to make new advances on identifying, formulating and solving existing open problems by combining insights from the different academic fields. Thus, we continued working on themes, initiated at the previous Dagstuhl Seminars, related to the representation of queuing in dynamic models and the identification of new requirements and new open problems arising from future forms of mobility and travellers' behaviour, for example, systems of autonomous vehicles, and coordinated/shared routing.

Again, the seminar was a big success. Beside forster collaboration of the last seminars of this series, the seminar stimulates new and very fruitful ones. We got laudatory feedback from many participants which is also reflected in the survey conducted by Dagstuhl.

6.28 The Constraint Satisfaction Problem: Complexity and Approximability

Organizers: Martin Grohe, Venkatesan Guruswami, Dániel Marx, and Stanislav Živný

Seminar No. 22201

Date: May 15–20, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.5.112

Participants: Kristina Asimi, Per Austrin, Libor Barto, Manuel Bodirsky, Bertalan Bodor, Andrei A. Bulatov, Silvia Butti, Lorenzo Ciardo, Victor Dalmau, Dmitry Feichtner-Kozlov, Martin Grohe, Venkatesan Guruswami, Johan Hastad, Mark R. Jerrum, Amirhossein Kazeminia, Michael Kompatscher, Pravesh K. Kothari, Marcin Kozik, Andrei Krokhin, Barnaby Martin, Dániel Marx, Antoine Mottet, TomášNagy, Tamio-Vesa Nakajima, Daniel Neuen, Joanna Ochremiak, Jakub Oprsal, Michael Pinsker, Aaron Potechin, Jakub Rydval, Madhur Tulsiani, Caterina Viola, Albert Vucaj, Uli Wagner, Dmitriy Zhuk, Stanislav Živný

Remote Participants: Albert Atserias, Amey Bhangale, Joshua Brakensiek, Hubie Chen, Florian Frick, Pasin Manurangsi, Dana Moshkovitz, Aditya Potukuchi, Akbar Rafiey, Sai Sandeep

The *constraint satisfaction problem*, or CSP in short, provides a unifying framework in which it is possible to express, in a natural way, a wide variety of computational problems dealing with mappings and assignments, including satisfiability, graph/hypergraph colorability, and systems of equations. The CSP framework originated around 40 years ago independently in artificial intelligence, database theory, and graph theory, under three different guises, and it was realised only in the late 1990s that these are in fact different faces of the same fundamental problem. Nowadays, the CSP is extensively used in theoretical computer science, being a mathematical object offering a good balance between generality and structure that provides an excellent laboratory both for classification methods and for algorithmic techniques, while in AI and more applied areas of computer science this framework is widely regarded as a versatile and efficient way of modelling and solving a variety of real-world problems, such as planning and scheduling, software verification and natural language comprehension, to name just a few. An instance of CSP consists of a set of variables, a set of values for the variables, and a set of constraints that restrict the combinations of values that certain subsets of variables may take. Given such an instance, the possible questions include (a) deciding whether there is an assignment of values to the variables so that every constraint is satisfied, or optimising such assignments in various ways, or (b) finding an assignment satisfying as many constraints as possible. There are many important modifications and extensions of this basic framework, e.g., those that deal with counting assignments or involve soft or global constraints.

Constraint satisfaction has always played a central role in computational complexity theory; appropriate versions of CSPs are classical complete problems for most standard complexity classes. CSPs constitute a rich and yet sufficiently manageable class of problems to give a good perspective on general computational phenomena. For instance, they help to understand which mathematical properties make a problem tractable (in a wide sense, e.g., polynomial-time solvable or non-trivially approximable, fixed-parameter tractable or definable in a weak logic). One of the most striking features of this study is the variety of different branches of mathematics (including universal algebra and logic, combinatorics and graph theory, probability theory and mathematical programming) that are used to achieve deep insights into CSP, and this Dagstuhl Seminar aims to contribute towards further synergy in the area.

After about 15 years of intense research activity and hugely impressive progress, the culmination of the algebraic-approach to fixed-template CSPs was the resolution of the Feder-Vardi conjecture independently by Bulatov and Zhuk in 2017. While some fundamental questions (such as a fine-grained understanding of tractable CSPs) remain open, new research directions on generalizations of CSPs started emerging. The fixed-template promise CSP (PCSP) is among the most promising new directions of research motivated by better understanding computational hardness or tractability. PCSPs are a vast generalization of CSPs where each predicate has a strong and a weak form and given a CSP instance, the objective is to distinguish if the strong form can be satisfied vs. even the weak form cannot be satisfied. A prime and well-known example is the approximate graph coloring problem: distinguish k-colorable graphs from graphs that are not even ℓ-colorable, for some fixed $k < \ell$. The main topic of this seminar is PCSPs, a highly ambitious research direction with

intriguing connections to both old open problems (such as the approximate graph coloring problem) and new research directions (such as generalizations of submodularity, a key concept in optimization).

The recent flurry of activity on the topic of the seminar is witnessed by five previous Dagstuhl Seminars, titled "Complexity of constraints" (06401) and "The CSP: complexity and approximability" (09441, 12541, 15301, 18231), that were held in 2006, 2009, 2012, 2015, and 2018 respectively. This seminar was a follow-up to the 2009, 2012, 2015, and 2018 seminars. Indeed, the exchange of ideas at the 2009, 2012, 2015, and 2018 seminars has led to ambitious new research projects and to establishing regular communication channels. There is clearly the potential for further systematic interaction that will keep on cross-fertilising the areas and opening new research directions. The 2022 seminar brought together 46 researchers from different highly advanced areas of constraint satisfaction and involved many specialists who use universal-algebraic, combinatorial, geometric, and probabilistic techniques to study CSP-related algorithmic problems. 10 of the participants attended remotely owing to the global COVID pandemic. The participants presented, in 24 talks, their recent results on a number of important questions concerning the topic of the seminar. One particular feature of this seminar is a significant increase in the number of talks involving multiple subareas and approaches within its research direction – a definite sign of the growing synergy, which is one of the main goals of this series of seminars.

Concluding remarks and future plans. The seminar was well received as witnessed by the high rate of accepted invitations and the great degree of involvement by the participants. Because of a multitude of impressive results reported during the seminar and active discussions between researchers with different expertise areas, the organisers regard this seminar as a great success. With steadily increasing interactions between such researchers, we foresee another seminar focusing on the interplay between different approaches to studying the complexity and approximability of the CSP. Finally, the organisers wish to express their gratitude to the Scientific Directors of the Dagstuhl Centre for their support of the seminar.

Description of the Topics of the Seminar

With the resolution of the Feder-Vardi conjecture on finite-domain CSPs by Bulatov and Zhuk in 2017, the field has moved on to more generalizations of finite-domain decision CSPs. One of the main emerging areas is that of Promise CSPs (PCSPs), which are a huge generalization of CSPs.

Promise CSPs The study of PCSPs is about approximability of perfectly satisfiable instances. This exciting area been the main theme of this workshop, with about half of the talks on PCSPs.

The workshop started with a 2-hour tutorial on the basics of PCSPs.

- In the first part, Opršal talked about basic tools used in the complexity analysis of PCSPs, namely minions and free structures.
- In the second part, Opršal talked about an application of topology in the study of complexity of graph colorings.

There were three topologists attending the seminar and one of them gave a talk.

- Kozlov gave a talk on uses of topology in theoretical computer science and homomorphism questions in particular.

There were 9 contributed talks on recent results on PCSPs.

- Barto talked about his work on "baby PCP", which is a combinatorial variant of a (weaker version of) the PCP theorem.
- Brakensiek gave two talks: one on minions related to SDP relaxations for PCSPs, and one on very recent results on robust solvability of certain PCSPs.
- Butti gave a talk about Sherali-Adams relaxations for valued PCSPs.
- Ciardo gave a talk about Sherali-Adams relaxations for PCSPs, and more generally about studying hierarchies of convex relaxations via tensorisation.
- Dalmau gave a talk about his work on a condition that PCSPs solvable by local consistency algorithms have to satisfy.
- Kompatscher gave an overview of existing results on finite tractability of PCSPs.
- Mottet gave a talk about his very recent work on first-order definable PCSPs and a relationship to PCSPs solved by local consistency algorithms.
- Nagajima gave a talk on the complexity of linearly-ordered colorings.

Approximability of CSPs The second main theme of the seminar was recent progress in the area of approximability of CSPs. This part of the programme started with a 2-hour tutorial:

- Tulsiani gave a tutorial on how to approximate CSPs via the Sum of Squares hierarchy of SDP relaxations, including related work on high-dimensional expanders.

Three participants gave talks on approximability of CSPs.

- Bhangale gave a talk on approximability of solvable linear equations over non-Abelian groups.
- Kothari presented a refutation algorithm for unsatisfiable instances of SAT and CSP.
- Potechin talked about an algorithm for Max NAE-SAT with clauses of lengths 3 and 5.

Infinite-domain CSPs The area of infinite-domain CSPs has seen some exciting progress in last few years with new techniques being developed.

- Bodirsky talked about methods for transferring complexity classification result between different classes of infinite-domain CSPs.
- Nagy talked about the complexity of CSPs over random hypergraphs.
- Pinsker gave an overview of recent developments around the dichotomy conjecture for infinite-domain CSPs.

Generalisations of CSPs There are several interesting research directions related to CSPs, including the number of solutions, CSPs with a bound on the number of occurrences for each variable, quantified CSPs, a generalisation to the ideal membership problem, etc. The seminar featured 5 talks on these topics.

- Austrin presented an algorithmic lowerbound for perfect matching on random graphs, a canonical example of an edge CSP.
- Bulatov gave a talk the ideal membership problem.

- Chen presented his results on property testing of homomorphism inadmissibility.
- Kazeminia presented a classification of counting graph homomorphism modulo 2.
- Zhuk presented a complexity classification of quantified CSPs.

6.29 Anticipatory Human-Machine Interaction

Organizers: Jelmer Borst, Andreas Bulling, Cleotilde Gonzalez, and Nele Russwinkel

Seminar No. 22202

Date: May 15–20, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.5.131

Participants: Henny Admoni, Elisabeth André, Mihai Bâce, Leslie Blaha, Jelmer Borst, Matteo Bortoletto, André Brechmann, Andreas Bulling, Cristina Conati, Emmanuelle Dietz, Steffen Frey, Cleotilde Gonzalez, Chenxu Hao, Susanne Hindennach, Christian P. Janssen, Samuel Kaski, Frank Keller, Oliver W. Klaproth, Alvitta Ottley, Nele Russwinkel, Andreea Ioana Sburlea, Florian Strohm, Niels Taatgen

There is growing interest in interdisciplinary understanding of the anticipatory processes regarding other people's actions and also one's own actions. This surge in interest is especially important given that anticipatory interaction is needed in the growing area of intelligent systems that work closely together with human partners. Decades of research on human-machine interaction (HMI) have resulted in significant advances in theories, tools, and methods to facilitate, support, and enhance interactions between humans and computing systems. Despite the fundamental importance of HMI for our information society and numerous advances towards making interactions with machines more human-like, current systems still fall short in one core human ability – Theory of Mind (ToM). ToM allows us to attribute mental states to others and anticipate their actions; and is thus essential for us to interact naturally, effortlessly, and seamlessly.

ToM shapes how we interact with each other and is most easily observable in physical tasks, such as moving a table together. In this scenario, we rely on ToM abilities to attribute intentions to others and in turn, continuously adapt our own behaviour to accommodate the intentions of others, resulting in seamless collaboration. ToM begins to develop in early childhood. Even small children who are not able to develop the full cognitive ability of ToM can anticipate others' intentions using simple mental models of tasks and following familiar sequences of actions (e.g., Fiebich, 2018).

Deficits in ToM are closely linked to developmental disorders, such as autism, and current machines are similarly mind-blind. That is, they fail to sense users' attention and predict their intentions, and therefore lack the ability to anticipate and pro-actively adapt to users' actions. This limits machines to operating after the fact, i.e., to merely react to user input. This fundamentally limits the naturalness, efficiency, and user experience of current interactions. Imagine how difficult it would be to move a table with a robot: as the robot cannot anticipate easy and hard actions for the user (e.g., the table is heavy and needs to be rested for a moment, a corner has to be navigated for which the user has to switch her grip, etc.), it would not be a very helpful teammate. We believe anticipatory HMI has significant potential to bring systems that possess ToM to a new, exciting level of development. However, the building blocks required for anticipatory HMI are still at a very early research stage. This is in stark contrast to the large body of work on (Computational) Theory of Mind in the cognitive and neurosciences and the variety of potential applications in which artificial anticipatory behaviour could have a revolutionary and paradigm-changing influence.

With this seminar we have made a first step in bridging this gap and have discussed theoretical foundations, key research challenges and opportunities, new computational methods, as well as applications and use cases of anticipatory HMI. Research on anticipatory HMI is inherently interdisciplinary and draws from a number of fields, most notably human-computer interaction, machine learning, computer vision, computational neuroscience, computational cognitive science, privacy and security, as well as context-aware computing. Consequently, in this seminar, we have brought together junior and established researchers within these communities for the first time to explore the scientific foundations of anticipatory HMI.

Three perspectives on human-machine anticipation

In the seminar we have discussed three perspectives on human-machine anticipation.

Anticipating the intentions, beliefs, mental states and next actions of the user. This perspective focuses on the challenges, as well as possible solutions, for machine anticipation of user intentions, goals, motivations, and behaviors.

This includes the underlying question of which mental states need to be modeled and anticipated in the first place. This might be different depending on the purpose of the collaboration or on the specific circumstances.

Anticipating the outcome of the machine's own actions. The ability of a machine to anticipate its own actions and outcomes is important, especially for interaction in teams and for learning purposes. Artificial agents need to have a concept of the outcome of their own actions, as that will naturally affect the collaboration with the human or other agents. In addition, in case the expected outcome of an action is not reached, they have to learn how they can achieve the expected outcome in an alternative manner.

Anticipating the outcome of collaborative work and human-machine teaming. The third perspective builds on the previous two. In human-machine teaming, a machine needs to anticipate the human partners' actions, identify problems, and develop ideas on how the partners in a team can be supported. We will discuss how artificial agents can engage in effective teamwork with humans and which specific abilities are required. We will also discuss whether we can construct artificial agents that behave indistinguishable from human agents – and whether this is even desirable.

Underlying concepts and methods

In addition to the three different perspectives on anticipatory behaviour, we have discussed the underlying concepts and computational methods.

On the conceptual side, we have discussed definitions and developed a common language. This included, for example, concepts such as cognitive state vs. mental states (which are not always concisely defined across disciplines). Furthermore, to support our cross-disciplined, multi-context approach, we have collected different forms and definitions of "anticipation". A related but separate concept that we have discussed is what different disciplines understand by ToM and what related terms that should be differentiated. We found that there are considerable differences of opinion on this that we formulated and exchanged.

On the methodological side, we have discussed the ways in which different domains and disciplines have their own methods, approaches, and challenges. Subsequently, groups have exchanged and discussed these collections to gain a deeper understanding of available methodological approaches. We have also talked about current approaches in modelling and cognitive science, and exchange ideas with those of us from other disciplines. Which approaches support the development of better applications? More general is the question how user behaviour can be measured. What data can or should be collected (CV, natural language, neurobiological measures, behavior,...), how can we process it most effectively, and how can we incorporate it into cognitive or computational models of the other users? Finally, we have discussed what the best way is to integrate and fuse different sources of information about the user, and how this can be integrated with the current context.

6.30 Exponential Analysis: Theoretical Progress and Technological Innovation

Organizers: Annie Cuyt, Wen-shin Lee, and Gerlind Plonka-Hoch

Seminar No. 22221

Date: May 29–June 3, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.5.170

Participants: Dmitry Batenkov, Bernhard Beckermann, Robert Beinert, Wen-Yang Chu, Annie Cuyt, Dirk de Villiers, Jürgen Gerhard, Johan Gielis, Mark Giesbrecht, Mariya Ishteva, Ferre Knaepkens, George Labahn, Wen-shin Lee, David Li, Ridalise Louw, Ana C. Matos, Hrushikesh N. Mhaskar, Jesús Ortega Almirón, Gerlind Plonka-Hoch, Daniel Potts, Jürgen Prestin, Tomas Sauer, Richard G. Spencer, Hanna Veselovska, Rina-Mari Weideman

For the analysis and representation of stationary signals and images the conventional Fourier- and wavelet-based methods are particularly appropriate. However, in many areas in science and engineering we are faced with the problem to interpret digital signals and images which are not band-limited and have a non-stationary behaviour. Frequently, there are even further obstacles. The acquisition of signal or image measurements may be very expensive and therefore limited. In other applications, measurement sets are huge but contaminated by noise. Examples of the above are encountered in magnetic resonance imaging, infrared microscopy, fluorescence-lifetime imaging microscopy (FLIM), the analysis of seismic signals in geophysics, radar imaging (SAR/ISAR), tissue ageing models, vibration analysis, direction of arrival (DOA) detection, texture classification, radio frequency identification (RFID), non-destructive testing, satellite navigation, time series analysis, echolocation, induction motor diagnostics (MCSA), to name just a few.

Within the last few years, research on Prony-based methods has been intensified, as they offer an alternative to the compressed sensing approach. One essential advantage of the Prony method is that it does not need randomly collected measurements but works with deterministic sampling based on a sampling scheme which is adapted to the nonlinear signal model. At the same time, the Prony approach does not suffer the well-known curse of dimensionality in the multivariate case.

This Dagstuhl Seminar brought together a number of researchers from different areas in mathematics, engineering and industry. Topics included new mathematical insights and efficient numerical algorithms for problems based on exponential analysis as well as applications of exponential analysis models in engineering and life sciences. During this Seminar, the participants presented their newest results and discussed several open problems and applications from different perspectives.

The talks in the workshop partially had the character of survey talks and were arranged in different main research topics, as new mathematical theory (Day 1 and Day 5), new computational approaches (Day 2) and new results in applications in engineering (Day 3) and life sciences (Day 4).

On the first day, the talks focussed on the **mathematical theory of exponential analysis**. B. Beckermann and A. Matos emphasized the close connection between exponential analysis and rational approximation, which leads to new application areas as density reconstruction in an equilibrium problem in logarithmic potential theory and improved rational approximation of Markov functions. As shown in the talk by T. Sauer, there is also a close connection between multi-exponential analysis and continued fractions.

In the afternoon, the participants started four smaller **thematic discussion groups** to discuss new challenges and open problems throughout the week. These discussion groups particularly focussed on theory, more efficient computational algorithms and applications in engineering and biosciences. The discussions included general observations on the further development of exponential analysis and its ties to other subjects as well as further specific approaches for application of the theory to application problems in radar imaging or MRI. Still, we are far away from a full understanding of the relations between the different methods for the stable reconstruction of a parametric signal model, as well as from a systematic construction of Fourier analytic methods for the improved analysis of non-stationary signals.

In practice, Prony-based methods often suffer from a bad conditioning of the involved structured matrices and some extra effort is required to reliably execute the corresponding algorithms. Among the more successful implementations, the ESPRIT method, the Matrix-Pencil method, the approximate Prony method, and validated exponential analysis are established. The problem statement is also closely related to rational approximation theory and the structured low-rank approximation of

structured matrices. These connections are still not completely understood and may lead to strongly improved reconstruction algorithms, able to treat more general sampling sets and deliver super-resolution results.

The second day was devoted to the problem of **efficient computations** for reconstruction or approximation of functions using multi-exponential analysis. First numerical methods like MUSIC are already implemented in the computer algebra system Maplesoft, as presented in the talk by J. Gerhard. The advantage of the application of Maplesoft is that it can process the data within a desired very high precision and therefore successfully handle these reconstruction problems which are known to be ill-conditioned. Another way to improve the numerical stability is the application of more sophisticated numerical approaches to structured matrices and the connection to rational approximation, such that the existing stable algorithms for rational approximation can be used, as presented in the talk of G. Plonka. The survey presentation by H. Mhaskar turned the attention to the application of neural networks to approximation problems, and D. Potts presented new efficient Fourier methods to compute the so-called ANOVA decomposition for approximation of multivariate functions.

In the afternoon a mini-course on newest features of Maplesoft was presented by J. Gerhard, with a special focus on the problem of how to connect MAPLESOFT with other Software as MATLAB etc. The remaining time was used for further scientific discussions in smaller groups.

On the third day, the talks surveyed different new **applications of exponential analysis in engineering**. The presentation of F. Knaepkens showed new approaches for direction of arrival (DOA) estimation, image denoising and inverse synthetic aperture radar. Chromatic aberration in large antenna systems have been studied by D. de Villiers. R.-M Weideman showed her results on antenna position estimation through sub-sampled exponential analysis of signals in the near-field. The talks by R. Beinert and J. Prestin showed applications concerning phase retrieval in optical diffraction tomography and detection of directional jumps in images.

The survey presentations on the fourth day focussed on **exponential models in life sciences**. J. Gielis showed in his talk, how generalized Möbius-Listing bodies (GML) can be employed for better modelling and understanding of certain dynamical processes in the natural sciences. D. Li explained recent progress of advanced time-resolved imaging techniques based on exponential analysis models and their applications in life sciences, for example to reveal biological processes at the molecular level. In a further talk, R.G. Spencer reported on newest developments in Magnetic Resonance Relaxometry and Macromolecular Mapping to achieve more accurate myelin quantification in the brain that permits the establishment of physiological correlations. The underlying reconstruction problem is a seriously ill-posed inverse problem.

The last day of the workshop was again devoted to further results in **mathematical theory of exponential analysis** and connections to other areas of mathematics. In the talk by D. Batenkow, the degree of ill-posedness of the parameter reconstruction problem based on the exponential sum model was studied in more detail. The degree of condition of the problem essentially depends on the distribution of the frequency parameters. H. Vesolovska discussed the problem of recovering an atomic measure on the unit 2-sphere $\mathbb{S}^2$ given finitely many moments with respect to spherical harmonics. A connection of exponential analysis to computer algebra problems was brought to our attention by M. Ishteva. She showed how the joint decomposition of a set of non-homogeneous polynomials can be computed using the canonical polyadic decomposition, and how this decomposition can be applied in nonlinear system identification.

This Dagstuhl Seminar has been an important milestone for improved understanding of the large impact of exponential analysis tools in both theory and practice. During the thematic discussions in small groups in this meeting, several new aspects have been considered and several collaborations have been initiated or continued. Examples include new approaches for an improved modelling of antenna system frequency responses in radio frequency (Cuyt, De Villiers, Weideman) and for stabilised parameter estimation in exponential models using iterative factorizations of matrix pencils of Loewner matrices (Beckermann, Plonka-Hoch).

We mention that this seminar is related to Dagstuhl Seminar 15251 on "Sparse modelling and multi-exponential analysis" that took place in 2015. The discussions at the latter have led to many interesting collaborative projects, among which a funded Horizon-2020 RISE project (Research and Innovation Staff Exchange) with the acronym EXPOWER, standing for "Exponential analysis Empowering innovation" (grant agreement No 101008231).

It is our experience that these Dagstuhl Seminars are timely and seminal. Through the meetings new collaborations and new potential are unlocked. There is a clear need to further connect stakeholders from the new theoretical developments and the identified industrial applications, as is our objective here and in the future.

6.31 Radical Innovation and Design for Connected and Automated Vehicles

Organizers: Wendy Ju, Bastian Pfleging, and Andreas Riener

Seminar No. 22222

Date: May 29–June 3, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.5.188

Participants: Jonas Andersson, Regina Bernhaupt, Gary Burnett, Mark Colley, Debargha Dey, David Goedicke, Kai Holländer, Wendy Ju, Andreas Löcken, Nikolas Martelaro, Nathasha Merat, Alexander Meschtscherjakov, Lutz Morich, Bastian Pfleging, Andreas Riener, Virpi Hannele Roto, Shadan Sadeghian Borojeni, Ronald Schroeter, David Sirkin, Helena K. Strömberg

Automated driving will most likely be the next big change in individual mobility. Currently, research is still primarily concerned with technical challenges of Automated Driving Systems (ADS), such as sensing, data processing, communication, or steering. HCI researchers have already started to develop concepts for working, relaxing, or recreational activities (in privately owned) ADS as rather incremental innovations. There is, however, room to think about creative ways to use automated vehicles (AV) and connected technology towards the public interest beyond incremental changes. Current open questions are, amongst others:

- What are possible innovative and groundbreaking visions for future human transport concepts?
- How can we apply radical innovation and design and leverage AV technology to other applications and use cases?
- What types of research and design methods are able to contribute to radical designs, instead of incremental?

This Dagstuhl Seminar aimed to contribute to continued research that is able to challenge the current generative/evaluative research approach for automated driving systems against a radical innovations attempt. Similar to the invention of the car, which was not just an incremental improvement of a horse carriage, we wanted to question whether the current incremental research approach is appropriate and want to provoke novel opportunities. We further wanted to challenge whether a human-centered design approach is appropriate in the domain of connected, automated vehicles or if related innovation-centered approaches like design thinking are more constructive.

An integral part of the seminar was to reach the next level of product quality, to learn from each other, and to disseminate individual's experience for boosting subsequent research. In order to give the seminar participants something they can use directly from the seminar, we dedicated two blocks of time for teaching methodological knowledge in addition to the creative sessions, interactive panels and practical work. The goal was to give the participants the opportunity to try out new methods, such as brainwriting, bodystorming, focus groups, user enactment, WorldCafe, amongst others, that they would like to use in their own research in the future – thus a directly applicable result of the seminar. Details on the seminar's schedule are shown in Fig. 6.2.

To promote trans- and interdisciplinarity, we invited computer scientists/engineers, interaction designers, UI/UX designers, market and consumer psychologists as well as urban planners from industry and academia to discuss and design future mobility and vehicle concepts. We expected three types of results from this seminar – and will discuss at the end of the full report which contributions were created in the seminar and to which degree these results were achieved by the end of the seminar.

- **New visions for transportation and interaction**. The seminar focused on radical and innovative visions for the future of human transport and employed innovation-centered approaches to foster out-of-the-box thinking.
- **Novel applications for AV technology**. Perhaps the most significant contribution of the seminar should be novel applications for automated vehicle technology that will stretch the research and science community into new directions. We expect that in the coming 3 to 10 years, these new ideas should serve as inspiration for the research of the seminar attendees and, more broadly, the communities involved in designing automotive and mobility user interfaces.
- **Roadmap(s) for research**. Another goal is the elaboration of a roadmap to outline proposed research collaborations and recommend new funding mechanisms. Furthermore, the roadmap should lay out plans for disseminating results such that members of our community are well informed, and such that they can effectively interact with researchers and practitioners in related communities, such as human-computer interaction, human-factors, user experience, automotive engineering, psychology, and urban/traffic planning.

Dagstuhl Seminar 22222: Draft Program

7:30-8:45 Breakfast				
Introduction	Design Methods II	Recap Tuesday	Prototyping	Plotting outcomes II
Mid-morning break				
Radical Innovation Panel	Novel UsersPanel	Prototyping	Plotting outcomes	Summary
12:15-13:45 lunch				
Design Methods I	Propose Radical Innovation Initiatives	Group photo	Drafting documents, videos, prototypes	
15:30-16:00 coffee/cake				
Walk and Talk: Mobility of the Future	World Cafe	Hiking	Dagstuhl homework	
18:00-19:00 dinner		Hiking dinner	18:00-19:00 dinner	
Movie Night			Q&D video screening	

Fig. 6.2
Schedule of Dagstuhl Seminar 22222 (May 29 to June 3, 2022).

Perhaps the most tangible and easy-to-communicate outcome of the seminar for other interested parties are short videos/video prototypes of ideas created over the week in group work. All of the videos were professionalized in the aftermath of the seminar and submitted, along with one short paper each, to the ACM AutomotiveUI conference, where three of the videos were accepted in the video track and shown to the audience at the conference in Seoul (videos are available as supplementary material in the ACM DL).

References

1 Nikolas Martelaro, Debargha Dey, Gary Burnett, Helena K. Strömberg, Jonas Andersson, and Andreas Löcken. 2022. How to Manage Social Order in Shared Automated Vehicles. In Adjunct Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '22). Association for Computing Machinery, New York, NY, USA, 201 – 203. https://doi.org/10.1145/3544999.3550154

2 Kai Holländer, Lutz Morich, Natasha Merat, Gary Burnett, Virpi Roto, Wendy Ju, and David Sirkin. 2022. Radical Innovations in Future Mobility and Virtual Assistants. In Adjunct Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '22). Association for Computing Machinery, New York, NY, USA, 209 – 211. https://doi.org/10.1145/3544999.3551501

3 Regina Bernhaupt, Mark Colley, David Goedicke, Alexander Meschtscherjakov, Bastian Pfleging, Andreas Riener, and Shadan Sadeghian. 2022. A Critical Perspective on Radically Innovating Personal Mobility. In Adjunct Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '22). Association for Computing Machinery, New York, NY, USA, 215 – 218. https://doi.org/10.1145/3544999.3551689

6

6.32 Theories of Programming

Organizers: Amy Ko, Thomas D. LaToza, David C. Shepherd, and Dag Sjøberg

Seminar No. 22231

Date: June 6–10, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.6.1

Participants: Sebastian Baltes, Jonathan Bell, Moritz Beller, Gunnar Bergersen, Michael Coblenz, Scott Fleming, Thomas Fritz, Tudor Girba, Andrew Head, Robert Hirschfeld, Brittany Johnson-Matthews, Jun Kato, Amy Ko, Thomas D. LaToza, Sarah Lim, Justin Lubin, Gail C. Murphy, Hila Peleg, Lutz Prechelt, Francisco Servant, David C. Shepherd, Dag Sjøberg, Justin Smith, Emma Söderberg, Kathryn T. Stolee, Jeffrey Stylos, Benjamin Xie, Andreas Zeller

Mature scientific disciplines are characterized by their theories, synthesizing what is known about phenomena into forms which generate falsifiable predictions about the world. In computer science, the role of synthesizing ideas has largely been through formalisms that describe how programs compute. However, just as important are scientific theories about how programmers write these programs. For example, software engineering research has increasingly begun gathering data, through observations, surveys, interviews, and analysis of artifacts, about the nature of programming work and the challenges developers face, and evaluating novel programming tools through controlled experiments with software developers. Computer science education and human-computer interaction research has done similar work, but for people with different levels of experience and ages learning to write programs. But data from such empirical studies is often left isolated, rather than combined into useful theories which explain all of the empirical results. This lack of theory makes it harder to predict in which contexts programming languages, tools, and pedagogy will actually help people successfully write and learn to create software.

Computer science needs scientific theories that synthesize what we believe to be true about programming and offer falsifiable predictions. Whether or not a theory is ultimately found to be consistent with evidence or discarded, theories offer a clear statement about our current understanding, helping us in prioritizing studies, generalizing study results from individual empirical results to more general understanding of phenomena, and offering the ability to design tools in ways that are consistent with current knowledge.

Dagstuhl Seminar 22231 on *Theories of Programming* explored the creation and synthesis of scientific theories which describe the relationship between developers and code within programming and social activities. The seminar brought together researchers from software engineering, human-computer interaction, programming languages, and computer science education to exchange ideas about potential theories of programming. We identified and proposed theories that arose from many sources: untested but strongly-held beliefs, anecdotal observations, assumptions deeply embedded in the design of languages and tools, reviews of empirical evidence on programming, and applications of existing theories from psychology and related areas. Our aim was to bridge this gulf: formulating deeply-held beliefs into theories which are empirically testable and synthesizing empirical findings in ways that make predictions about programming tools and languages.

To achieve this aim, the seminar had three specific goals.

1. Bring together researchers with diverse expertise to find shared understanding.
2. Create a body of theories which make testable predictions about the effects of programming tools, languages, and pedagogy on developer behavior in specific contexts.
3. Propose future activities which can advance the use of theories, including identifying studies to conduct to test theories and ways to use theories to communicate research findings to industry.

During this seminar, a few short talks first reviewed the nature, creation, and use of theories as well as existing evidence about developer behavior during programming activities. The main activity of the seminar was working in small groups to sketch new theories of programming.

Seminar Overview

The seminar was divided into the following sessions across four days in June 2022:

- Tuesday: welcome, what is theory, describing theories, critiquing theories

- Wednesday: brainstorming unexplained programming phenomena, sketching theories, getting feedback on theories, and refining theories
- Thursday: presenting theory sketches, discussing ways of sharing theories, and skeptically examining whether developing theories of programming is really worth the time
- Friday: reflecting on takeaways and departure

The seminar was organized by Thomas Latoza, Amy J. Ko, Dag Sjøberg, David Shepherd, and Anita Sarma. Anita later had to drop out, leaving Thomas, Amy, Dag, and David as the four organizers who were able to attend.

What is theory?

The goal of this opening session was to find common ground on what theory was. To achieve this, each organizer gave short presentations related to theories.

Thomas identified how researchers used theories to generate falsifiable predictions about the world. He described common characteristics of theories as abstract, explanatory, relevant, and operationalizable. An example of a theory of programming he provided was how violating constraints cause defects or reduces code quality.

Amy described an interpretivist framing of theories, where theories were cultural and experience-based. Some theories were folk theories (e.g., code is magic, Not Invented Here, and spaces vs tabs for white space). Some theories were personal, such as programming as common sense machines and tinkering towards correctness. Other theories came from research communities such as ICSE. For example, a theory of programming is that we can copy and adapt code from another location in a program to fix bugs.

Dag drew guidelines between what was and was not a theory. He identified multiple examples of what were not theories: scientific laws were not theories because they were missing the "why;" trivial statements were also not theories. The building blocks of theories included constructs, propositions, explanations, and scope. Theories can help us explain surprising empirical results, while empirical results can help us support or refute certain theories. Finally, Dag noted how premature theorizing is likely to be wrong, but can still be useful.

David emphasized the importance of keeping theories practical. He defined a relationship from theorems to corollaries to examples and applications. He provided an example of using different representations in music for different use cases and users.

In open discussions and breakout groups, attendees identified additional nuances to theories. We noted how it is useful for theories to enable ease of communication or shared understanding. But by defining a vocabulary, theories can also limit the scope of explanation. We can also use theories to understand what we observe or to justify interventions. Finally, there was discussion about creating theories inductively, deductively, and/or abductively.

Common themes that arose from discussion include how theories are seldom used to justify the design of programming languages and tools, and how programming is a social endeavor and drawing upon social science research (e.g., psychology) can support theory building.

Expressing Theories using a Theory Template

The goal of this session was to try to express theories using a theory template developed by the organizers. While the goal of this template was to support the creation of new theories, attendees used it to describe existing theories for this session. Attendees broke into five groups to attempt to apply the theory template to the following existing theories of programming:

- Asking and answering questions [1]
- Program comprehension as fact finding [2]
- Leaky abstractions [3]
- Information hiding [4]
- Theory of programming instruction [5]

After considering feedback from attendees, organizers revised the theory template. The revised theory template's section headers and helper text are as follows:

1. *Theory's name*: Choose a name that is memorable, short, and descriptive.
2. *Summary*: In a few sentences, summarize the phenomena, constructs, relationships, and a concrete example, hypothesis, and study.
3. *Contributors*: Who has contributed to this theory? Add your name here.
4. *Phenomena*: What programming phenomena is your theory trying to explain? And in what scope (people, expertise, contexts, tools, etc.)? This description should just describe what is being explained, should not offer an explanation; that below. ("Programming" includes any and all interactions between people and code, in any context (e.g., software engineering, learning, play, productivity, science, and all of the activities involved in creating programs, including requirements, architecture, implementation, verification, monitoring, and more).
5. *Prior Work*: What prior work offers an explanation of this phenomena, or might help generate an explanation of this phenomena? For the purposes of the seminar, this does not need to be complete, but a complete description of this theory would have an extensive literature review covering theories that inspired this theory, as well as conflicting theories.
6. *Concepts*: Describe the key concepts of the theory and some concrete examples of them, building upon the phenomena above. These might be variables, processes, people, aspects of people, structures, contexts or other phenomena that are essential to the theory's account of the phenomena. Note: concepts should be descriptions of ideas that give some structure and precision to describing the phenomena, not operationalizations or measurements – those belong in example hypotheses and/or studies.
7. *Relationships and Mechanisms*: Using the constructs described above, explain the causality of how the phenomena works. What causes what and how? Provide a few concrete examples to illustrate the idea.
8. *Example Hypotheses*: What testable claims do the constructs, relationships, and mechanisms imply?
9. *Example Studies*: What are existing or envisioned example study methods that might investigate the hypotheses above? How might the concepts be operationalized and measured? Describe details about populations, samples, tasks, contexts, tools, observations. Remember that studies can involve many forms of observation and data, both qualitative and

quantitative and even design contributions. Studies do not have to be feasible to be proposed and can vary in scope, from single-study sized methods to long-term research agendas that might explore a theory over many years and many projects.

10. *Corollaries*: What follows from this theory, if true? Provide potential implications, concrete or otherwise.

Unexplained phenomenon

After spending the first day discussing what theories were and applying a theory template, the goal of the second day was to identify unexplained phenomena related to programming and apply theories to explain them. After an informal voting process, attendees created groups to develop theories around the following phenomena:

- Debugging
- Types
- Neurodiversity in programming
- Data programming
- Code examples
- Developer tools
- Learning effects from code analysis

Groups spent all of Wednesday developing theories by filling out the theory template and then getting feedback from members of other groups. They then iterated and created presentations. For descriptions of each presentation, see the talks' abstracts included in the full report.

Sharing Theories

On Thursday after the presentations, attendees had discussions about how to share theories of programming to broader audiences. Many ideas included written dissemination, such as publishing research, writing books, creating a wiki, adding to reviewer guidelines, creating a website, defining syllabi for reading groups, speaking on podcasts, and posting on social media sites. Other ideas featured opportunities for further interaction, such as workshops, special interest groups, demonstrations of theories for practitioners, and stickers/flair for engagement at conferences. Other ideas focused on incentive structures, such as creating a "best new theory" award at conferences.

Group-wide discussions about sharing theories identified some structural barriers and opportunities. A barrier to broader theory creation and/or use is that most computing researchers do not have much training in theories. Workshops, reading groups, or changes to undergraduate or graduate level coursework could help address this. Another structural barrier is that most conferences lack instructions about theory. Adding instructions in paper calls and reviewer instructions as well as "theory shepherds" could help address this systemic barrier.

Do we really need theories?

The final session for Thursday was critically reflective about whether programming actually required theories. Given this session occurred after lunch on the final full day, this session got silly. After splitting into groups to discuss, groups shared eclectic presentations to reflect their discussions:

- a colorful whiteboard diagram about pros and cons of theories (Fig. 6.3)
- Another whiteboard diagram about whether to use theory (Fig. 6.4)
- A list of bullets about challenges of making changes in publishing
- An humorous improv skit about a conference Q&A session on a theory paper.

Reflections on the week

The final session of this seminar asked attendees to reflect on the seminar as a whole. Attendees identified some high-level takeaways:

Attendees found theories useful for helping understand why things (e.g., languages or tools) do or do not work. They also found theories helpful for differentiating between how we think people work and how they actually work.

Attendees also felt that the engagement of computing researchers with theories of programming was often limited by the lack of interest and/or lack of expertise. Interdisciplinary research can help create the gestalt of expertise required to create theories of programming, but narrow conference and journal scopes often make this difficult. Specifically, many computing researchers lack expertise in empirical evaluations, making it difficult to develop rigorous evidence that is often foundational to theory building. Furthermore, much training in empirical evaluations focuses on lab settings, whereas most programming happens "in the wild."

Multiple attendees also felt that theories were more implicitly prevalent in computing research than was explicitly discussed. Some conversation focused on "lower case 't'" theories, or theories that we not fully formalized, but provided use and explanation. Many attendees felt that theories implicitly were existing in papers, but were unaware of explanations into this work.

A concluding consensus was that theories of programming have existed in the background. Through explicit engagement and discourse, this Dagstuhl Seminar could serve as a catalyst to augment existing theories and craft new ones.

References

1 Sillito, Jonathan, Gail C. Murphy, and Kris De Volder. "Asking and Answering Questions during a Programming Change Task." IEEE Transactions on Software Engineering 34, no. 4 (July 2008): 434–51. https://doi.org/10.1109/TSE.2008.26.

2 LaToza, Thomas D., David Garlan, James D. Herbsleb, and Brad A. Myers. "Program Comprehension as Fact Finding." In Proceedings of the the 6th Joint Meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on The Foundations of Software Engineering, 361–70. ESEC-FSE '07. New York, NY, USA: Association for Computing Machinery, 2007. https://doi.org/10.1145/1287624.1287675.

3 Spolsky, Joel. "The Law of Leaky Abstractions." Joel on Software, November 11, 2002. https://www.joelonsoftware.com/2002/11/11/the-law-of-leaky-abstractions/.

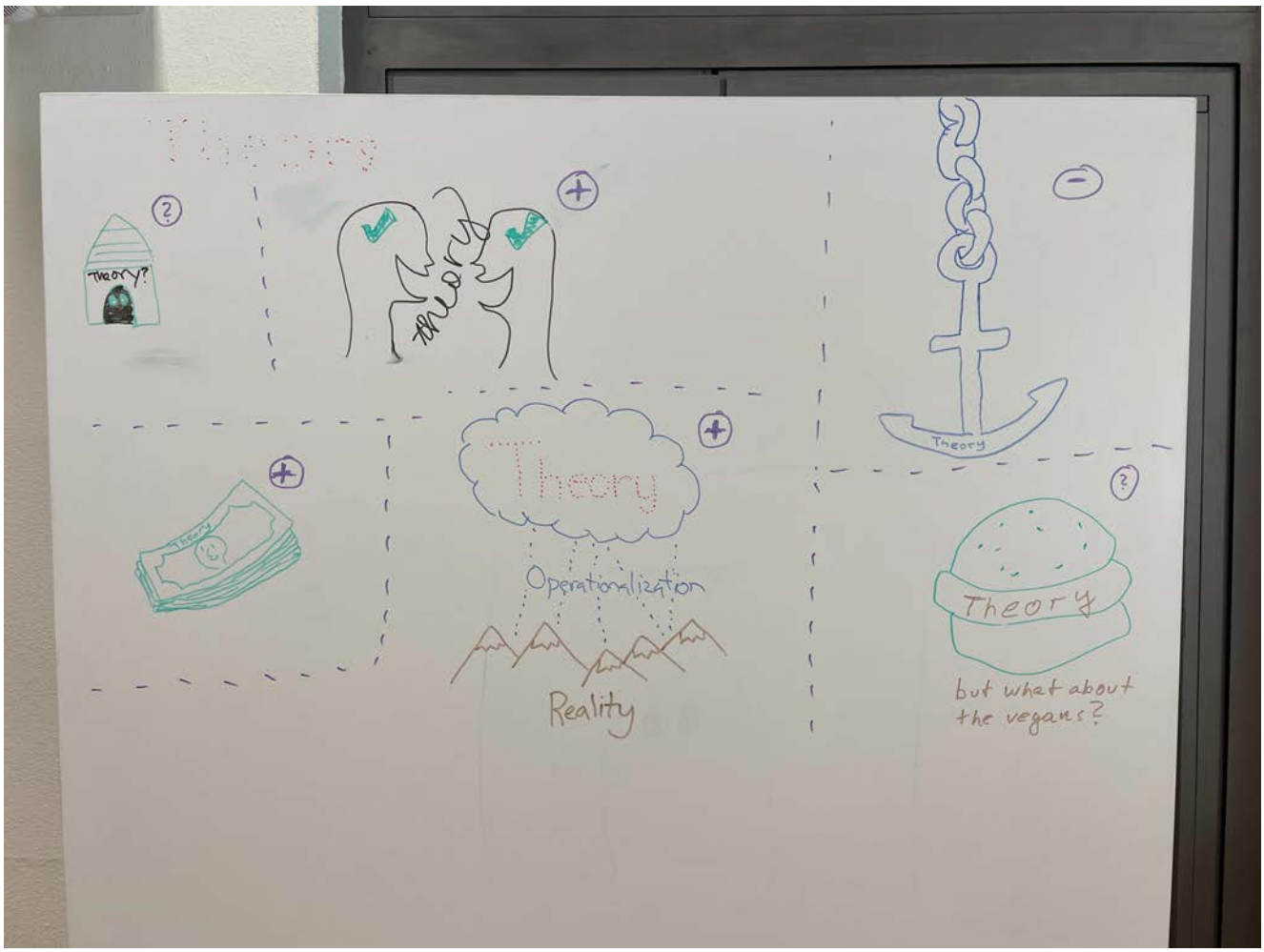


Fig. 6.3
Whiteboard sketches of the pros and cons of theories, as depicted by various diagrams.

4 Parnas, David L. "On the criteria to be used in decomposing systems into modules." Pioneers and their contributions to software engineering. Springer, Berlin, Heidelberg, 1972. 479–498.

5 Xie, Benjamin, Dastyni Loksa, Greg L. Nelson, Matthew J. Davidson, Dongsheng Dong, Harrison Kwik, Alex Hui Tan, Leanne Hwa, Min Li, and Amy J. Ko. "A Theory of Instruction for Introductory Programming Skills." Computer Science Education, January 25, 2019, 1–49. https://doi.org/10.1080/08993408.2019.1565235.

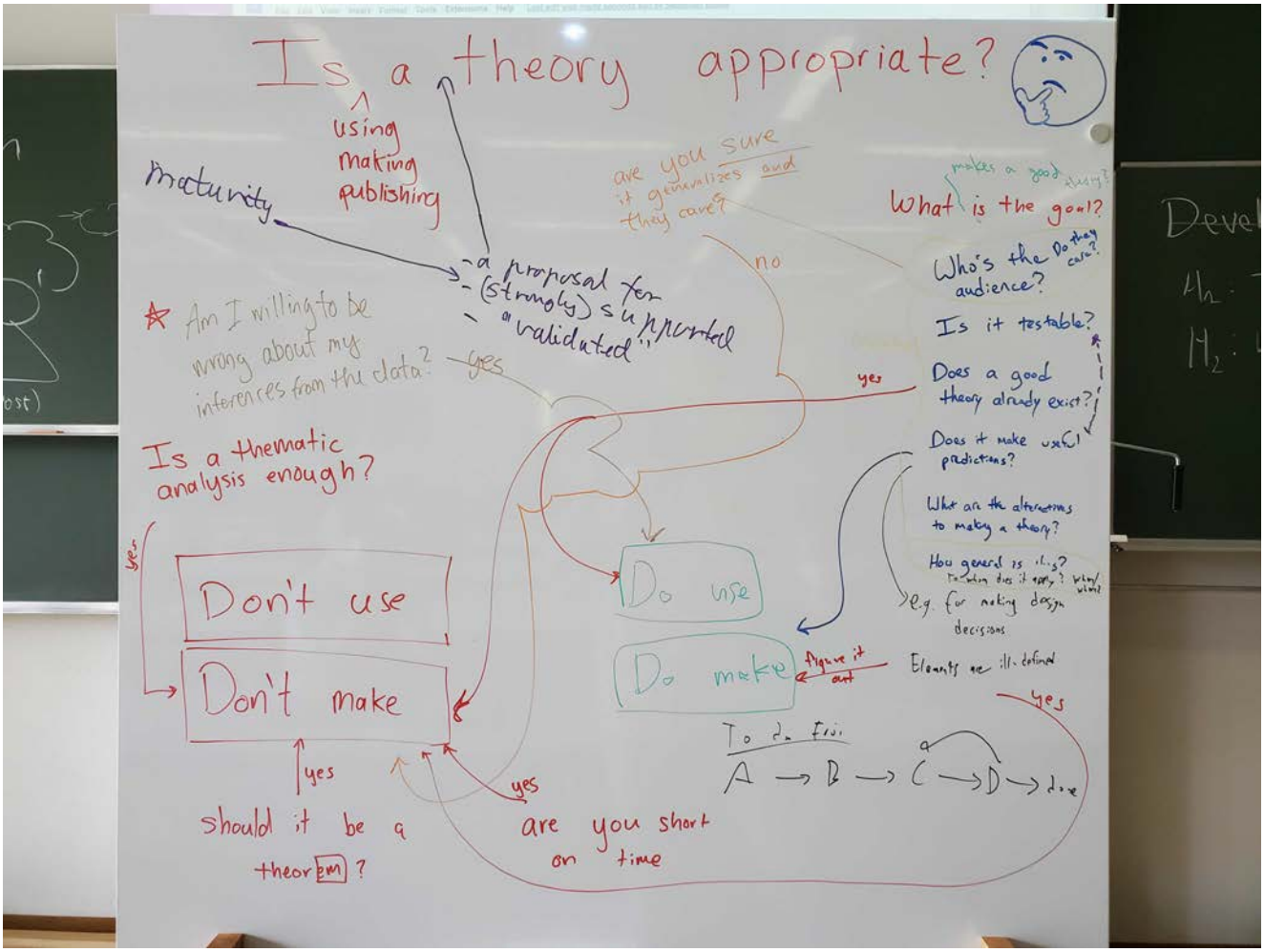


Fig. 6.4
Whiteboard sketch of flowchart considering whether a theory is appropriate.

6.33 Efficient and Equitable Natural Language Processing in the Age of Deep Learning

Organizers: Jesse Dodge, Iryna Gurevych, Roy Schwartz, and Emma Strubell

Seminar No. 22232

Date: June 6–10, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.6.14

Participants: Yuki Arase, Niranjan Balasubramanian, Leon Derczynski, Jesse Dodge, Jessica Forde, Jonathan Frankle, Iryna Gurevych, Michael Hassid, Kenneth Heafield, Sara Hooker, Alexander Koller, Ji-Ung Lee, Alexander Löser, Alexandra Sasha Luccioni, André F. T. Martins, Haritz Puerto, Colin Raffel, Nils Reimers, Leonardo Ribeiro, Anna Rogers, Andreas Rücklé, Roy Schwartz, Edwin Simpson, Noam Slonim, Noah A. Smith, Emma Strubell, Betty van Aken, Thomas Wolf

For this seminar, we brought together a diverse group of researchers and practitioners in NLP and adjacent fields to develop actionable policies, incentives and a joint strategy towards more efficient and equitable NLP. This Dagstuhl Seminar covered a range of related topics, which we summarize as follows.

Efficient NLP models A key method for mitigating the raised concerns is reducing costs by making models more efficient. We surveyed the different methods that exist for making NLP technology more efficient. We discussed their tradeoffs, prioritized them, and aimed to identify new opportunities to promote efficiency in NLP. During the seminar, we drafted a survey paper summarizing multiple methods for increasing the efficiency of NLP models. We aim to publish this work later this year.

Systemic issues We also addressed systemic issues in the field relating to the reporting of computational budgets in NLP research, and how we can use incentive structures such as the NLP Reproducibility Checklist [1] to motivate researchers throughout the field to improve reporting. We discussed the survey responses for the reproducibility checklist used at four major NLP conferences, and we plan to release a report of this data.

Equity of access A third topic of discussion was the equity of access to computational resources and state-of-the-art NLP technologies. Prior to the seminar, we conducted a survey of different stakeholders across the NLP community. During the seminar, we analyzed and discussed the results of this survey to better understand who is most affected and how, and developed informed strategies and policies to mitigate this inequity moving forward. We are currently working on a paper summarizing the results of this survey, which we hope to publish later this year.

Measuring efficiency and equity All of the above endeavors require establishing the right metrics and standards to measure our current status and progress towards efficiency and equity goals. We discussed multiple metrics and evaluation frameworks that capture the bigger picture of how different approaches compare in terms of energy efficiency not just in the research environment but in practice and over the entire ML model lifecycle (development, training and deployment), and that work under a wide range of computational budgets.

References

1 Jesse Dodge, Suchin Gururangan, Dallas Card, Roy Schwartz, Noah A. Smith: Show Your Work: Improved Reporting of Experimental Results. EMNLP/IJCNLP (1) 2019: 2185–2194

6.34 Human-Game AI Interaction

Organizers: Dan Ashlock, Setareh Maghsudi, Diego Perez Liebana, and Pieter Spronck

Seminar No. 22251

Date: June 19–24, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.6.28

Participants: Maren Awiszus, Cameron Browne, Duygu Cakmak, Alex J. Champandard, Guillaume Chanel, Michael Cook, Alena Denisova, Alexander Dockhorn, Manuel Eberhardinger, Jakob Foerster, Casper Harteveld, Amy K. Hoover, Ahmed Khalifa, Antonios Liapis, Daniele Loiacono, Simon M. Lucas, Setareh Maghsudi, Ana Matran-Fernandez, Paris Mavromoustakos Blom, Mark J. Nelson, Mirjam Palosaari Eladhari, Diego Perez Liebana, Mike Preuß, Lisa Rombout, Jacob Schrum, Pieter Spronck, Tommy Thompson, Remco Veltkamp, Vanessa Volz, Jichen Zhu

Over the past decades, artificial intelligence has evolved from esoteric techniques used mainly in computer science research to an integral and ever-growing part of the daily lives of most humans. People regularly interact with semi-intelligent machines during their daily lives, whether it is via smartphone applications, embedded systems in cars and household electronics, client support systems, or helpful technology installed on personal computers. People wish and expect systems to respond intelligently to their requests, and even to anticipate their actions. While improvements to the interaction between humans and intelligent systems in this respect have been made over the years, there is still a long way to go before these systems exhibit a level of understanding and intuition which can be expected from a human partner.

Human-computer interfaces (HCI) are a well-established scientific research domain. We noted that HCI research generally neglects the use of artificial intelligence as a integral part of an interface. Almost any person that uses computers can quickly recall multiple frustrating interactions with the current state of the art in artificial intelligence in interfaces. Since annoyance and apparent incompetence can derail the adoption of otherwise promising and potentially transformative technology, research into improving interfaces using AI is timely.

An AI assistant is broadly recognized as being a key factor in increasing human productivity, but it must be an AI assistant that the user either enjoys working with or that the user barely notices, not one that must be bludgeoned into useful behavior or constantly fought with. Perfection of assistants, companions, and even opponents that correctly anticipate and collaborate in the relatively controlled domain of games provides a smooth path to such developments in broader contexts.

We argue that virtual worlds, as found in computer games, are an ideal environment in which to experiment with the interaction between humans and artificial intelligence. There are at least three reasons for this. First, virtual worlds often approach the complexity of the "real world", while still being under the control of the researcher and completely observable. Second, the agents in virtual worlds are supposed to represent "real people" and are approached as such by the humans who "play" with the virtual world. Third, the potential interactions that players have with the virtual worlds are highly diverse and wide-ranging, which presents a substantial challenge for artificial intelligence to respond to in a reasonable fashion.

In recent years, the number of ways in which human players can interact with games have increased considerably. While ten years ago interaction was almost exclusively through mouse and keyboard or controllers, nowadays games can potentially respond to natural movements and facial expressions captured by a camera, to spoken language, to eye movements, and to signals captured by a variety of sensors. Brain-computer interface (BCI) technology has become more mainstream, offering possibilities for games to respond to a users' brain activity. Using VR technology, games can respond to movements of players in natural space. AI that can use all these interface elements to make game agents a natural and appreciated partner or opponent for humans can form the basis for advanced AI agents that interact with humans not only in games, but also in the real world.

The research area in which the seminar is rooted is the interaction between humans and game AI, aiming for natural and appropriate responses of computational agents in virtual worlds to human behavior, making use of both traditional interaction technology as well as modern sensor and interaction technology.

The research area lends itself for a wide range of research topics. For the preparation of this seminar, we proposed the following set of sub-topics:

- **Personalized Human-Game AI Interaction:** Humans have different backgrounds, interests, and goals. As such, there is no "one-size-fits-all" interference and interaction form.

Under this topic, we explore game adaptation as a type of automatic game design. The goal is to permit the AI to adapt the game environment to the player based on the observed features and received feedback. Instead of fully automatic game design, a sophisticated game design leaves scope for an AI to adapt to a broad variety of players. Such personalized adaptation could be extended to adaptation of the actual game interface – in games, usually complex interactions are possible, which novice players are not capable of employing. Therefore, automatically adapting the interface to the observed experience level of the player may be a valid approach to effective personalization.

- **Human-Game AI Interaction for People with Disabilities:** People with disabilities require special attention when designing interfaces, to mitigate adverse effects of disabilities, so that a suitable experience is ensured for everyone. Game AI can potentially help to diagnose disabilities, both physically and psychologically. There is also the potential for game AI to create awareness of issues faced by those with disabilities, by intelligently adapting the interface in such a way that the player experiences it as a person with disabilities would.
- **Multimodal Interfacing and Interaction:** Multimodal systems offer a flexible and efficient interaction environment that consists of several input/output possibilities including text, speech, and vision. How to effectively use these possibilities in game design is still an open problem. A compelling application of artificial intelligence is to rapidly learn which modes a given player finds natural and enjoyable. The type of interface a user is comfortable with is likely to cross boundaries between different applications, meaning that an "interface fingerprint" may be derivable that can be carried with the user, permitting the re-use of information gained.
- **Enhancing Human Creativity with Artificial Intelligence:** Computational Creativity is a field of AI where automatic AI systems design and create various forms of art, which may include images, drawings, poetry and music. In the broader sense, these systems create new content either completely by themselves, or with the human providing input at specific points. Research into this fusion of the creative skills of humans and AI systems would move the state of the art a step forward: from being inferior content creators the AI systems would become a tool for amplifying and augmenting the superior creative abilities of a human being, in a bi-directional collaboration process. AI systems should be able to learn from the human, anticipate what they intend to do, and understand the domain of discourse. They would provide advice on content creation and help when the user struggles with certain techniques or creative methods. By learning the skills of the human, AI systems would be able to propose alternatives that lie outside their expertise, allowing the humans to learn, refine and improve their capabilities. The users would experience a system that adapts to their skills, needs and pace, and becomes a personalized companion in their learning process.
- **Trustful and Reliable Human-Game AI Interaction:** We often observe that humans feel uncomfortable with AI recommendations. Moreover, mistakes made by humans are deemed more tolerable than those made by an AI. While there is no objective rationale for this difference, it is hard to justify the use of AI for humans by arguing that AI offers a lower mistake probability compared to humans. It is therefore imperative to find new ways to convince humans to interact with the game AI and to take its advice seriously. Moreover, it is crucial to minimize any effect that might harm such trust, regardless of its origin.
- **Information Flow in Human-Game AI Interaction:** A game AI must observe the human player and, in turn, provide players with information that they find helpful, valuable, or interesting. Even the most potentially helpful information is not actually helpful if the player cannot understand it or if it is not useful to their particular style of play. The flow of information is particularly important between the human player and an AI companion. Reliable metrics that ascertain if the human uses information offered by the AI, that check if the AI fails to provide information that the human tries to find in other ways, and assessment of defects in the human's play that suggest which information is needed, are potential goals of research in this area.
- **Believable Human-Game AI Interaction:** In the last decade, contests have been held at several conferences where human judges voted on the "humanity" of both human game players and AI players in an effort to score the ability of the AI players to behave in a plausibly human manner. Attempts to make AIs interact in a way that is indistinguishable from human interaction are a natural way to structure research into human-game AI interaction. We note that the believability of game AI often suffers because it fails to recognize that it misunderstands the human player, or that the human player misunderstands the AI. How to recognize misunderstanding, followed by how to correct for misunderstanding, are important steps in making game AI more believable.
- **Ethics of Human-Game AI Interaction:** Several of the aforementioned research directions rely heavily on big data analysis. Acquiring such a massive amount of data is a challenging task. Perfect anonymization is hard to achieve, and often undesirable as multiple parties are involved in data collection and integration. To what extent is it ethical to collect personal interaction information? Are there ethical restrictions to the extent to which an AI is allowed to analyze a player's personality and demographics? These questions need answering even if a player gives permission to collect and use such data.
- **Novel Forms of Interaction and Interfaces in Game AI:** New technology gives rise to new possibilities in game interaction and interfacing. While developers often try to restrict themselves to small adaptations in tried-and-true forms of interaction, it makes sense to consider the interaction possibilities originating with novel technology, such as virtual reality and brain-computer interfacing. Beyond those, there may be ways for humans and AI to interact with each other that has not yet been imagined, or which can benefit from re-imagining. Player-AI interaction can be implemented in many forms, such as (1) cuing a player with environmental information from music to decor, (2) influencing a player by adjusting game elements such as local architecture, opponents, and rewards, and (3) making a player respond to the social tone of non-player characters. Such alternate forms of player-AI interaction warrant investigation.

This seminar was organized around workgroups, which worked in teams and topics proposed by the participants of the seminar in the areas outlined above. These workgroups were accompanied by plenary sessions for group formation, topic debate and discussions of the deliberation of each group. Workgroups were dynamic, so participants could move between them, and new groups were formed during the week. A Discord server was setup for coordination and announcements, and it was

also used by the different groups for document and link sharing. This also has the benefit of providing a place for discussions after the seminar, easing the communication and further work among the members of each workgroup.

It is worthwhile mentioning the work carried out during the invitation process. Due to the COVID crisis, the changes in the political landscape, and the war in Ukraine, many declined the invitation, and many participants dropped out after originally having accepted the invitation. Thus, multiple rounds of invitations were run until two weeks before the seminar. We invited close to 100 people, the full list of invitations having a high diversity (a 50% male-female split, about half invitations for 'junior' people, and invitees hailing from all continents – including South America and Africa, which are usually highly underrepresented). In the end, just over 30 participants attended the seminar (out of the 45 possible). Size-wise this was a slight disappointment. We were fortunate, however, that those that did attend were highly enthusiastic and highly knowledgeable about the topics covered, which made the seminar a great success.

6.35 Visualization Empowerment: How to Teach and Learn Data Visualization

Organizers: Benjamin Bach, Sheelagh Carpendale, Uta Hinrichs, and Samuel Huron

Seminar No. 22261

Date: June 26–July 1, 2022, Dagstuhl Perspectives Workshop 22261 | Dagstuhl Perspectives Workshop
Full report – DOI: 10.4230/DagRep.12.6.83

Participants: Jan Aerts, Fearn Bishop, Peter C.-H. Cheng, Alexandra Diehl, Jason Dykes, Sarah Hayes, Uta Hinrichs, Trevor Hogan, Christoph Huber, Samuel Huron, Mandy Keck, Christoph Kinkeldey, Søren Knudsen, Doris Kosminsky, Tatiana Losev, Areti Manataki, Isabel Meirelles, Luiz Morais, Till Nagel, Rebecca Noonan, Georgia Panagiotidou, Laura Pelchmann, Fateme Rajabiyazdi, Jonathan C. Roberts, Christina Stoiber, Yagoda Walny, Wesley J. Willett

Remote Participants: Wolfgang Aigner, Benjamin Bach, Magdalena Boucher, Robert S. Laramee, Andrew Manches, Alison Powell, Mashael Alkadi, Mine Çetinkaya-Rundel, Andy Kirk, Fanny Chevalier, Nathalie Henry Riche, Dietmar Offenhuber, Sheelagh Carpendale, Marti Hearst, Charles Perin, Emeline Brulé

This seminar set out to discuss timely issues and approaches to teaching and education in data visualization. The topic is of growing importance in a world where more and more content is being shared through online news and social media. Our mission as researchers, practitioners and educators in data visualization is to assure quality education for everyone engaging with visualization; this ranges from visualization designers, data scientists, school teachers, journalists, working professionals, students, as well as general public audiences. Teaching visualization is tricky for a range of reasons:

- Data visalization is a skill that is only slowly starting to make its way into school curricula (at least in some countries);
- While the **range of visualization tools** available makes it easy for almost anyone to create visualizations regardless of their technical background, it can be overwhelming to know where to start and to navigate this ever-growing and changing tool landscape;
- Data visualization is a highly interdisciplinary field, influenced and moved forward by psychology, cognitive science, design, computer science, data science, art, and many more disciplines. As a result, **learning objectives and teaching practices greatly vary**;
- There is currently **no defined agreement on the learning goals and criteria** for visualization literacy. For example, what defines a beginner, intermediate, professional in data visualization? What aspects of data visualization should be taught at different levels? And: how can we assess visualization skills?;
- From a learner perspective, the motivation to pick up visualization as a skill is broad: some people "just" want to use a specific tool to get things done quickly, others pursue a design approach (no coding language required), others want to build systems for visualization (Computer Sciences), others go on and become educators or researchers;
- Visualization is important in many domains and knowledge and specific solutions might be specific to these domains,

rather than valid universally (e.g., color choices, symbolic conventions, level of interactivity);
- There are a lot of **tacit knowledge and skills** involved in visualization which can be difficult to pin down and transform into learning activities.

In order to discuss these challenges and how to navigate them, we invited participants from academia and industry, including senior and junior thinkers.

Participants & Seminar Format Given the highly interdisciplinary field of data visualization and visualization literacy, participants covered a range of expertises including the fields of design, computer science, human-computer interaction, education, graphics, and cognitive psychology. The 5-day seminar was run in a hybrid format with 28 participants joining us at Schloss Dagstuhl in-person, 7 participants joining us online synchronously from Europe, and 6 participants joining us asynchronously from North America. Two organizers were on-site at Schloss Dagstuhl, while two joined the seminar remotely (one synchronously and one asynchronously). One of the online organizers led the asynchronous North America group from Canada. All seminar participants (synchronous and asynchronous) met for a daily debriefing session at 5pm local (Dagstuhl) time to share their progress and discussions. The synchronous remote participants (Europe) joined different local discussion groups through online calls, which did work out surprisingly well – *special thanks to the Dagstuhl technical team for the amazing help with the hybrid setup*.

Seminar Structure & Activities The seminar followed an open-ended approach with respect to the possible outcomes, to allow discussion topics to emerge and develop, based on participants' expertise and interests. Discussions were sparked by brief talks and visualization activities led by selected participants.

The **seminar talks** included presentations on visualization teaching and learning with children, a syntactic analytical framework for visualization, engaging new students with visualization, using forums to engage students with visualization content, how to approach and streamline large-scale assessment of university students' visualization projects, as well as an overview over a book project from a past Dagstuhl Seminar (find the complete list and abstracts of talks in Section 5 of the full report).

From a practical end, the **visualization activities** invited seminar participants to actively engage in and experience a number of visualization teaching methods and techniques (see Section 6 of the full report for more details). One activity invited participants to sketch their relation to the seminar topic in order to introduce participants to each other and to start immersing them into the seminar topic. Another activity asked participants to analyze a given visualization systematically. In one activity, we classified existing visualization activities that were submitted by participants prior to the seminar. Another activity took a speculative approach to visualization, inspiring critical visualization scenarios and designs through a card game.

There was ample time to discuss topics of interests through **breakout groups** which focused on topics related to
- Teaching methods and taxonomies for educational activities;
- Teaching creativity and criticality for visualization;
- Data physicalization and how corresponding methods can be used for education and engagement;
- Practical approaches to teaching visualization and the politics involved in teaching visualization;
- Approaches to visualization teaching and creation inspired by improvisation in the arts, and eventually;
- Grand challenges in visualization education.

From an organizer perspective, the seminar was a great success. All participants – both on-site and online – were extremely engaged, and we obtained very positive feedback. Participants appreciated the creative and open-ended nature of this seminar that invited for sharing and reflection of practices from different disciplines and perspectives. The seminar produced a long list of outcomes ranging from paper outlines and book projects, to collecting teaching manifestos and taxonomies, to grant projects and platforms for sharing teaching tools and resources. The plan emerged to establish a reoccurring international symposium around visualization education as part of the IEEE VIS conference, the largest annual conference on visualization with over 1000 participants. The individual working groups will move their individual goals forwards after the seminar. As organizers, we will coordinate between groups and support each of the projects as best as we can, e.g., through regular check-ins with the workgroup leaders as well as townhouse meetings with all Dagstuhl participants, e.g., once a semester. We all believe strongly that this Dagstuhl Seminar – the first formal event on visualization education besides smaller conference workshops – has created a strong momentum for visualization empowerment and education, and we are looking forward to sharing our outcomes on a dedicated website soon.

6.36 Human-Centered Artificial Intelligence

Organizers: Virginia Dignum, Wendy E. Mackay, John Shawe-Taylor, and Frank van Harmelen

Seminar No. 22262

Date: June 26–July 1, 2022 | Dagstuhl Perpectives Workshop
Full report – DOI: 10.4230/DagRep.12.6.112

Participants: Michel Beaudouin-Lafon, Mehul Bhatt, Stefan Buijsman, Mohamed Chetouani, Ulises Cortés, Adam Dahlgren Lindström, Emmanuelle Dietz, Marko Grobelnik, Alípio Jorge, Antonis C. Kakas, Samuel Kaski, Janin Koch, Helena Lindgren, Paul Lukowicz, Wendy E. Mackay, Ozlem Ozmen Garibay, Janet Rafner, Laura Sartori, John Shawe-Taylor, Jacob Sherson, Marija Slavkovik, Philipp Slusallek, Frank van Harmelen, Katharina A. Zweig

This workshop brought together 22 participants with a diverse background from AI, Robotics, HCI, Ubiquitous Computing, Business and Sociology, from across Europe and North America and laid the groundwork for a manifesto on Hybrid Human-centered AI systems.

Informed by currently ongoing large initiaves such as the EU-funded Humane AI Net, the Dutch Hybrid Intelligence Center, the Danish Centre for Hybrid Intelligence, and the OECD AI policy framework, four pillars of the manifesto emerged: (a) Collaboration and Cooperation, (b) Control & Adaptivity, (c) Transparency & Explainability, and (d) Societal dimensions. For each of these pillars, the workshop resulted in (i) key terminology, (ii) key research questions, (iii) metrics and methodologies, and (iv) benchmarks and challenges.

The above resulted in a solid framework for the Hybrid Human-Centered AI manifesto to be written by the partcipants in the months following the Dagstuhl Perspectives Workshop.

6.37 Algorithms for Participatory Democracy

Organizers: Markus Brill, Jiehua Chen, Andreas Darmann, and David Pennock

Seminar No. 22271

Date: July 3–8, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.7.1

Participants: Ben Abramowitz, Haris Aziz, Dorothea Baumeister, Gerdus Benadè, Robert Bredereck, Markus Brill, Alfonso Cevallos, Andreas Darmann, Théo Delemazure, Edith Elkind, Piotr Faliszewski, Rupert Freeman, Ashish Goel, Paul Gölz, Umberto Grandi, Matthias Greger, Davide Grossi, Jobst Heitzig, Ayumi Igarashi, Anson Kahng, Christian Klamler, Sonja Kraiczy, Martin Lackner, Jérôme Lang, Nicholas Mattei, Reshef Meir, Stefan Napel, Arianna Novaro, David Pennock, Dominik Peters, Marcus Pivato, Clemens Puppe, Daniel Reeves, Simon Rey, Ulrike Schmidt-Kraepelin, Ehud Shapiro, Sofia Simola, Nimrod Talmon, Jens Witkowski, William S. Zwicker

Participatory democracy aims at a broad and direct participation of citizens in policy decision making, enabling a large fraction of citizens to propose ideas, debate issues, and vote on decisions. Modern-day participatory democracy processes entail several kinds of *algorithmic* challenges. This seminar focused on the algorithms underlying three types of participatory democracy systems: (1) online decision-making platforms for governments and organizations (such as *LiquidFeedback* or *decidim*), (2) participatory budgeting processes that enable citizens to directly and collectively decide how to spend tax dollars, and (3) collective decision-making systems involving currency. We also had dedicated sessions discussing algorithmic challenges related to *liquid democracy* and the relation between participatory democracy and blockchain technology. Working groups have been initiated discussing partial participation in voting mechanisms, the use of currency in social choice problems, participatory budgeting, and the impact of computational social choice.

The technical program was complemented by a demo session in which Jobst Heitzig demonstrated *vodle* (http://www.vodle.it) and Daniel Reeves demonstrated Beeminder (https://www.beeminder.com/) and other decision-making tools. Moreover, we organized a panel discussion on *The Past, Present, and Future of Computational Social Choice*, moderated by Piotr Faliszewski. In this panel discussion, Haris Aziz, Edith Elkind, Jérôme Lang, and Bill Zwicker gave their perspectives on the development of the field of COMSOC.

The organizers thank all participants for their interesting ideas and viewpoints presented in talks, discussions, and informal meetings. Moreover, we would like to express our gratitude towards Schloss Dagstuhl and its staff for all the support before and during the seminar, which contributed to making this seminar a successful one.

6.38 Eat-IT: Towards Understanding Interactive Technology and Food

Organizers: Masahiko Inami, Sohyeong Kim, Florian 'Floyd' Mueller, and Marianna Obrist

Seminar No. 22272

Date: July 3–8, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.7.19

Participants: Ferran Altarriba Bertran, Sahej Claire, Christopher Dawes, Jialin Deng, Masahiko Inami, Kyung seo Jung, Sohyeong Kim, Nadejda Krasteva, Kai Kunze, Neharika Makam, Florian 'Floyd' Mueller, Marianna Obrist, Harald Reiterer, Matti Schwalk, Jürgen Steimle

Remote Participants: Eleonora Ceccaldi, Simon Henning, Rohit Khot, Maurizio Mancini, Patrizia Marti, Nandini Pasumarthy, Yan Wang

In July 2022, 21 researchers and academics from Europe, Australasia and the USA gathered for a week to discuss the future of the coming together of food and information technology (IT), shortly called eat-IT.

Eating is a basic human need, and there is a growing interest in the field of Human-Computer Interaction (HCI) in designing new interactive food experiences, for example, to promote healthier food practices (e.g., [3,4,12,17]), to make eating a more enjoyable experience (e.g., [6, 18, 19, 21, 22]), and to design multisensory eating experiences (e.g., [13,15,16]). Theoretical work around the design of interactive food also emerged, for example, Grimes and Harper [5] proposed that a new view on human-food interactions (HFI) is required and introduced the concept of "celebratory technology" that emphasizes the positive aspects of eating in everyday life. Computational technology can make a significant contribution towards such celebratory technology, for example, Khot et al. [7] presented a system called TastyBeats that instead of presenting physical activity data on a screen, it offers users personalised sports drinks where the quantity and flavour is based on the amount of exercise a user has done. In a similar vein, EdiPulse [8] was introduced as a system that creates activity treats (chocolate creations) using a food printer. The shape and quality of the prints were based on the person's physical activities on that day, allowing for personal and shared reflections through consuming chocolate instead of looking at graphs on a screen. Computer science and in particular the information visualization community can, therefore, regard food (and drinks) as a medium to make data more approachable for people, communicating complex information in an easy-to-digest format [10]. Parametric design approaches have also influenced the way food is produced. For example, Wang et al. [20] developed the concept of shape-changing and programmable food that transforms during the cooking process. Through a material-based interaction design approach, the authors demonstrated the transformation of 2D into 3D food (i.e., pasta). They proposed these transformations for new dining experiences that can surprise users, but this can also be used for outer space, where food comes as a flat design and only transforms into a 3D form through the cooking process.

Furthermore, technological advancements in acoustic levitation have led to the design of taste-delivery technology that transports and manipulates food in mid-air [18], allowing for novel interactions between diners and food that is of interest to HCI researchers as it allows to study augmented food experiences without the use of cutlery. This work further extends taste stimulation towards a multisensory experience of levitating food due to the integration of smell, directional sound, lights, and touch [19]. Furthermore, robots are now in use to serve ice cream to the general public [2]. Lastly, laser-cutters have already been used to embed data into cookies through the engraving of QR codes [14]. Taken together, these examples suggest that computing technology can play a major role in the way we engage with food, in particular, there is a realization that technology can both facilitate instrumental benefits in regards to food (such as improved health through better food choices) as well as experiential benefits (such as enriched social experiences). In summary, computational technology has the potential to influence how people experience eating.

However, this notion of what we call "interactive food" also raises significant concerns. Will computing technology distract from the pure pleasures of eating? Will people accept meals that are optimized through data-driven approaches? Will people enjoy food that is served by robots? Will people understand and act on

data that is embedded in food? Questions such as these and, of course, their answers are important for the future of the field, and the seminar tried to investigate them.

The seminar was based on the belief that computer scientists, designers, developers, researchers, chefs, restaurateurs, producers, canteen managers, etc. can learn from each other to positively influence the future of interactive food. Working together allows for the identification of new opportunities the field offers but will also highlight the challenges that the community will need to overcome. In particular, it is still unknown what theory to use to design such computational systems in which the interaction is very multisensorial, contrasting the traditional mouse, keyboard and screen interactions. Furthermore, it is unclear how interacting with food is benefiting from, and also challenged by, our mostly three-times-a-day engagement with it (breakfast, lunch and dinner), again different to our interactions with mobile phones that occur at any time.

Furthermore, how do we create and evaluate interactions with computationally-augmented food that needs preparation time, again very different from our usually immediate interactions with interactive technology? What interaction design theory can guide us in answering these questions in order to extend computer science also to include food interactions? If such theoretical questions could be answered, as a flow-on effect, more insights could be generated on how to evaluate the success of such interactions. The result will be not only more engaging eating experiences, but also the potential to influence when and how and what people eat. This can have major health implications, possibly address major issues such as overeating that results in obesity and then a higher risk of diabetes, heart disease, stroke, bone and joint diseases, sleep apnea, cancer, and overall reduced life expectancy and quality of life [1].

Interrogating such topics is important, as otherwise industry advances will drive the field forward that can easily dismiss or oversee negative consequences when it comes to combining computational technology and food. It is imperative to get ahead of the curve and steer the field in the right direction through an interdisciplinary approach involving a set of experts brought together through the seminar.

Although there is an increasing number of systems emerging, there is limited knowledge about how to design them in a structured way, evaluate their effectiveness and associated user experiences as well as how to derive theory from them to confirm, extend or reject an existing theory. The seminar, therefore, examined these in order to drive a more positive future around interactive food.

Understanding the role of computational technology in this area is a way to make a positive contribution and guide the field in a positive way. There are a couple of areas of imminent importance, and we highlight these here:

- With advances in ubiquitous sensing systems, such as wearables, personal data becomes available in abundance. Such personal data can be embedded into food in order to either communicate it to users in engaging ways or as a way to personalize the food, such as when presenting meals that contain only those calories previously expended. If users understand such data visualizations or want to eat size-controlled portions based on personal data is an ongoing question. Issues of privacy and sharing of such data with chefs and kitchens are also open questions to be investigated.

- With advances in persuasive technology (as already utilized in the form of mobile apps that aim to persuade people to eat more healthy food), new opportunities arise to combine multiple sensor data from an IoT infrastructure to develop more persuasive systems. How people adhere to such approaches and change their behaviour to the better is still an underdeveloped area that needs to be investigated.

- Big data already used individually by producers, manufacturers and kitchens will increasingly converge, allowing to monitor and influence the supply chain from the farm to the diner's plate. This can help to optimize the sourcing of local produce, reducing the environmental impact through reduced transport distances and a reduction of food waste. How to make sense of such data and use machine learning and other AI advances to utilize this data, so it makes a difference to every part of this supply chain, is an important area for future work.

- Advanced sensor systems can now sense eating actions, such as through jaw movement [9]. By using machine learning, we can now gain an increased understanding of how people eat. This can inform the design of interactive systems that help people make better future eating choices. For example, people might stop overeating if a system could tell them that their stomach will produce a "full" feeling earlier than the 20 minutes it usually takes.

- With the advances of mixed-reality systems like VR headsets and augmented reality on mobile phones, new opportunities arise on how to augment food. For example, prior work has shown that people who perceive cookies through the use of augmented reality to be bigger than they actually are will change how much they eat [11]. There is therefore a significant opportunity to employ mixed-reality to change and offer new opportunities on what and how we eat. How to draw from and incorporate multimodal interaction design theory already established in computer science is an open question for the community to investigate.

- Robotic systems allow to prepare and serve food in novel and interesting ways, for example, robotic arms can already be purchased to be installed in personal kitchens and robots already serve ice cream in public ice cream shops. How to design such interactions so that they are engaging and safe, while still considering the joy and benefit from being engaged in cooking activities, is an interesting area for future work.

- Multisensory integration research has allowed to better understand how humans integrate sensory information to produce a unitary experience of the external world. It is reasonable to expect that technology will keep advancing and sensory delivery will become more accurate. In addition, our understanding of the human senses and perception will become more precise through large scale data from HCI and integration research. As such, there is the potential to systematize our definition of multisensory experiences, through adaptive, computational design. This is exciting but at the same time carries big questions on the implications of multisensory experiences as well as our responsibility when developing them.

The seminar began with talks by all attendees, in which they presented their work in the area and what they thought the biggest challenges are from their perspective the field is facing. After the presentations concluded, no more slides were used for the remainder of the week, with all activities being conducted either as a townhall meeting or in breakout groups. This was supplemented by optional morning and evening activities, such as jogging, beach volleyball, foosball, or cycling.

The structure of the seminar was based around theory, design and their intersection.

References

1 The world health report 2010.

2 Home – niska retail retail, Oct 2021.

3 Rob Comber, Eva Ganglbauer, Jaz Hee-jeong Choi, Jettie Hoonhout, Yvonne Rogers, Kenton O'hara, and Julie Maitland. Food and interaction design: designing for food in everyday life. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pages 2767–2770. 2012.

4 Rob Comber, Jaz Hee jeong Choi, Jettie Hoonhout, and Kenton O'Hara. Designing for human – food interaction: An introduction to the special issue on "food and interaction design". *International Journal of Human-Computer Studies*, 72(2):181–184, 2014.

5 Andrea Grimes and Richard Harper. Celebratory technology: new directions for food research in hci. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 467–476, 2008.

6 Rohit Ashok Khot, Deepti Aggarwal, Ryan Pennings, Larissa Hjorth, and Florian 'Floyd' Mueller. Edipulse: investigating a playful approach to self-monitoring through 3d printed chocolate treats. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 6593–6607, 2017.

7 Rohit Ashok Khot, Jeewon Lee, Deepti Aggarwal, Larissa Hjorth, and Florian 'Floyd' Mueller. Tastybeats: Designing palatable representations of physical activity. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2933–2942, 2015.

8 Rohit Ashok Khot, Ryan Pennings, and Florian 'Floyd' Mueller. Edipulse: supporting physical activity with chocolate printed messages. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pages 1391–1396, 2015.

9 Naoya Koizumi, Hidekazu Tanaka, Yuji Uema, and Masahiko Inami. Chewing jockey: augmented food texture by using sound based on the cross-modal effect. In *Proceedings of the 8th international conference on advances in computer entertainment technology*, pages 1–4, 2011.

10 Florian 'Floyd' Mueller, Tim Dwyer, Sarah Goodwin, Kim Marriott, Jialin Deng, Han D. Phan, Jionghao Lin, Kun-Ting Chen, Yan Wang, and Rohit Ashok Khot. Data as delight: Eating data. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2021.

11 Takuji Narumi, Yuki Ban, Takashi Kajinami, Tomohiro Tanikawa, and Michitaka Hirose. Augmented perception of satiety: controlling food consumption by changing apparent size of food with augmented reality. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 109–118, 2012.

12 Marianna Obrist, Rob Comber, Sriram Subramanian, Betina Piqueras-Fiszman, Carlos Velasco, and Charles Spence. Temporal, affective, and embodied characteristics of taste experiences: A framework for design. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 2853–2862, 2014.

13 Marianna Obrist, Yunwen Tu, Lining Yao, and Carlos Velasco. Space food experiences: Designing passenger's eating experiences for future space travel scenarios. *Frontiers in Computer Science*, 1, 2019.

14 Johannes Schoning, Yvonne Rogers, and Antonio Kruger. Digitally enhanced food. *IEEE pervasive computing*, 11(3):4–6, 2012.

15 Carlos Velasco and Marianna Obrist. Multisensory experiences: A primer. *Frontiers in Computer Science*, 3, 2021.

16 Carlos Velasco, Marianna Obrist, Gijs Huisman, Anton Nijholt, Charles Spence, Kosuke Motoki, and Takuji Narumi. Editorial: Perspectives on multisensory human-food interaction. *Frontiers in Computer Science*, 3, 2021.

17 Carlos Velasco, Marianna Obrist, Olivia Petit, and Charles Spence. Multisensory technology for flavor augmentation: a mini review. *Frontiers in psychology*, 9:26, 2018.

18 Chi Thanh Vi, Asier Marzo, Damien Ablart, Gianluca Memoli, Sriram Subramanian, Bruce Drinkwater, and Marianna Obrist. Tastyfloats: A contactless food delivery system. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, pages 161–170, 2017.

19 Chi Thanh Vi, Asier Marzo, Gianluca Memoli, Emanuela Maggioni, Damien Ablart, Martin Yeomans, and Marianna Obrist. Levisense: A platform for the multisensory integration in levitating food and insights into its effect on flavour perception. *International Journal of Human-Computer Studies*, 139:102428, 2020.

20 Wen Wang, Lining Yao, Teng Zhang, Chin-Yi Cheng, Daniel Levine, and Hiroshi Ishii. Transformative appetite: shape-changing food transforms from 2d to 3d by water interaction through cooking. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pages 6123–6132, 2017.

21 Yan Wang, Zhuying Li, Robert Jarvis, Rohit Ashok Khot, and Florian 'Floyd' Mueller. iscream! towards the design of playful gustosonic experiences with ice cream. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–4, 2019.

22 Zhuying Li, Rakesh Patibanda, Felix Brandmueller, Wei Wang, Kyle Berean, Stefan Greuter, and Florian 'Floyd' Mueller. The guts game: towards designing ingestible games. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, pages 271–283, 2018.

6.39 Security of Machine Learning

Organizers: Battista Biggio, Nicholas Carlini, Pavel Laskov, and Konrad Rieck

Seminar No. 22281

Date: July 10–15, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.7.41

Participants: Hyrum Anderson, Giovanni Apruzzese, Verena Battis, Battista Biggio, Wieland Brendel, Nicholas Carlini, Antonio Emanuele Cinà, Thorsten Eisenhofer, Asja Fischer, Marc Fischer, David Freeman, Kathrin Grosse, Pavel Laskov, Aikaterini Mitrokotsa, Seyed Mohsen Moosavi-Dezfooli, Nicola Paoletti, Giancarlo Pellegrino, Fabio Pierazzi, Maura Pintor, Konrad Rieck, Kevin Alejandro Roundy, Lea Schönherr, Vitaly Shmatikov, Nedim Srndic

Overview Modern technologies based on machine learning, including deep neural networks trained on massive amounts of labeled data, have reported impressive performances on a variety of application domains. These range from classical pattern recognition tasks, for example, speech and object recognition for self-driving cars and robots, to more recent cybersecurity tasks, such as attack and malware detection. Despite the unprecedented success of technologies based on machine learning, it has been shown that they suffer from vulnerabilities and data leaks. For example, several machine-learning algorithms can be easily fooled by adversarial examples, that is, carefully-perturbed input samples aimed to thwart a correct prediction. These insecurities pose a severe threat in a variety of applications: the object recognition systems used by robots and self-driving cars can be misled into seeing things that are not there, audio signals can be modified to confound automated speech-to-text transcriptions, and personal data may be extracted from learning models of medical diagnosis systems.

In response to these threats, the research community has investigated various defensive methods that can be used to strengthen current machine learning approaches. Evasion attacks can be mitigated by the use of robust optimization and game-theoretical learning frameworks, to explicitly account for the presence of adversarial data manipulations during the learning process. Rejection or explicit detection of adversarial attacks also provides an interesting research direction to mitigate this threat. Poisoning attacks can be countered by applying robust learning algorithms that natively account for the presence of poisoning samples in the training data as well as by using ad-hoc data-sanitization techniques. Nevertheless, most of the proposed defenses are based on heuristics and lack formal guarantees about their performance when deployed in the real world.

Another related issue is that it becomes increasingly hard to understand whether a complex system learns meaningful patterns from data or just spurious correlations. To facilitate trust in predictions of learning systems, the explainability of machine learning becomes a highly desirable property. Despite recent progress in development of explanation techniques for machine learning, understanding how such explanations can be used to assess the security properties of learning algorithms still remains an open and challenging problem.

This Dagstuhl Seminar aimed to bring together researchers from a diverse set of backgrounds to discuss research directions that could lead to the scientific foundation for the security of machine learning.

Goal of the Seminar The seminar focused on four main themes of discussion, consistent with the research directions reported above:

- Attacks against machine learning: What attacks are most likely to be seen in practice? How do existing attacks fail to meet those requirements? In what other domains (i.e., not images) will attacks be seen?
- Defenses for machine learning: Can machine learning be secure in all settings? What threat models are most likely to occur in practice? Can defenses be designed to be practically useful in these settings?
- Foundations of secure learning: Can we formalize "adversarial robustness"? How should theoretical foundations of security of machine learning be built? What kind of theoretical guarantees can be expected and how do they differ from traditional theoretical instruments of machine learning?
- Explainability of machine learning: What is the relationship between attacks and explanations? Can interpretation be trusted?

Overall Organization and Schedule The seminar intended to combine the advantages of conventional conference formats with the peculiarities and specific traditions of Dagstuhl events. The seminar activities were scheduled as follows:

Schedule	Activities
Day 1	Workshop presentation, short self-introductions by participants, one keynote presentation
Day 2	One keynote presentation on participant results, contributed presentations
Day 3	One keynote presentation on negative results, organization of working groups, one breakout session
Day 4	One breakout session, social event
Day 5	Keynote presentation, reporting from breakout sessions, summary of results

6.40 Current and Future Challenges in Knowledge Representation and Reasoning

Organizers: James P. Delgrande, Birte Glimm, Thomas Meyer, Miroslaw Truszczynski, and Frank Wolter

Seminar No. 22282

Date: July 10–15, 2022 | Dagstuhl Perspectives Workshop
Full report – DOI: 10.4230/DagRep.12.7.62

Participants: Michael Beetz, Meghyn Bienvenu, Piero Andrea Bonatti, Diego Calvanese, Anthony Cohn, James P. Delgrande, Marc Denecker, Thomas Eiter, Esra Erdem, Birte Glimm, Andreas Herzig, Ian Horrocks, Jean Christoph Jung, Sebastien Konieczny, Gerhard Lakemeyer, Thomas Meyer, Magdalena Ortiz, Ana Ozaki, Milene Santos Teixeira, Torsten Schaub, Steven Schockaert, Michael Thielscher, Francesca Toni, Renata Wassermann, Frank Wolter

Knowledge Representation and Reasoning (KR) is the field of Artificial Intelligence (AI) that deals with explicit, declarative representations of knowledge along with inference procedures for deriving further, implicit information from these symbolic representations. Research in KR as a mature area of AI is commonly taken as being marked by an Artificial Intelligence Journal Special Issue on Nonmonotonic Reasoning in 1980. In 1989 the Principles of Knowledge Representation and Reasoning Conference was founded, providing a dedicated, specialised forum for research in the area. While KR is one of the oldest and best-established areas of AI, it has continued to grow and thrive over the years. Most of the original research areas have evolved significantly, and have matured from the discovery and exploration of foundations, to the development and analysis of systems for emerging or established applications. Yet other areas, such as argumentation, arose much more recently, and are now thriving areas of KR.

While progress in KR has been steady and often impressive, it has not kept pace with the recent significant successes in AI in the use of statistical techniques and machine learning (ML). As a result, much of the work in AI, and much of the public perception of AI, centres on machine learning and on statistical applications. Nonetheless, we take it as given that KR is a vital, essential area of AI, and that research and development in KR remains necessary. Indeed, despite the unquestionable successes in machine learning and statistical techniques, limitations of these approaches are now emerging that, we believe, can only be overcome with advances in KR. Indicative of this is the recent interest in "Explainable AI", which requires a reference to declarative structures and reasoning over such structures. Furthermore, and in common with the majority opinion in AI, cognitive science, and philosophy, we take it as given that symbolic, declarative representations of knowledge are essential for any ultimate, general theory of intelligence.

For all of these reasons, a reassessment of the area of Knowledge Representation was a very timely undertaking of the Dagstuhl Perspectives Workshop 22282 "Current and Future Challenges in Knowledge Representation and Reasoning". During the seminar, the participants assessed the current state of KR along with future trends and developments. A questionnaire, which had been earlier distributed to the participants, helped in this assessment. Alltogether, the seminar served as a basis for developing an innovative agenda for the next 10–20 years of KR research. Key findings are measures to support a synergistic relationship with other subareas of the rapidly-changing field of AI and of computer science as a whole, e.g, through tutorials at the major KR conference, through new conference tracks and updated reviewing guidelines. The seminar further identified research areas for emphasis, assessed prospects for practical application of techniques, and considered how KR may address limitations of statistical techniques and machine learning.

The program comprised invited talks, panel discussions, working groups, and general dicussions. While the invited talks were agreed upon beforehand, the topics of the working groups (apart from Day 1) were decided interactively with all participants to allow for flexibility and reacting to the talks and the triggered discussions. Day 1 started with a short welcome and participant introduction session, followed by an assessment of *the past and present of KR* in the form of two invited talks by Anthony Cohn and Thomas Eiter. The remainder of Day 1 was dedicated to presenting the questionnaire results, which also prepared for the first working group on rethinking the call for papers (CfP) for the main KR conference, which not only served as rethinking the CfP, but also steered the working groups into thinking about the definition of KR as an area. The day closed with a report from the four working groups and indeed identified changes for the CfP, but also for the track structure and the recruitment and instructions for reviewers.

Day 2 focussed on the relationships of KR with four neighboring areas. For each sub-areas we began with a short invited talk (20 min) followed by a commentary (5 min), also invited, and a short general discussion (5 min). The function of the commentator was to look at the area from a different angle or give another perspective to avoid a too personal or narrow a perspective. The four talks addressed "KR and AI" (Ian Horrocks, commentator: Sébastien Konieczny), "KR and ML" (Francesca Toni, commentator: Ana Ozaki), "KR and Information Systems" (Diego Calvanese, commentator: Meghyn Bienvenu), and "KR and Robotics" (Gerhard Lakemeyer, commentator: Michael Beetz). Working groups on research challenges for these subareas concluded the day.

The third day began with a short talk on "Handling Uncertainty" (Jean Christoph Jung), for initiating a panel discussion on this topic. The morning concluded with a continuation of the working groups on sub-areas of KR from the previous day. The afternoon was dedicated to hiking and biking in smaller groups.

Day 4 started with short invited talks on "Applications of KR" (Esra Erdem, Thorsten Schaub, Michael Tielscher). The remainder of the day was dedicated to working groups on assessing the state of the art in sub-areas of KR and to expanding KR. For this latter group, we discussed the fact that geographically KR is stronger in Europe than in other parts of the world. As well, we considered how to attract new talent and how to reach out to disadvantaged groups, along with thinking of new forms of events such as hybrid conferences or virtual seminar series.

The final day of the seminar looked at strengthening the interaction between sub-areas of KR and wrapped up with statements of the participants regarding their personal impressions and "take-home" messages. This has, for example, already led to the creation of a novel KR discussion channel (on a Discord server). Key findings include that KR applications are very important to make the field visible and that applications are to be made more visible, e.g., through a journal special issue. Another outcome includes measures to reach out to other areas of AI, in particular machine learning and statistical techniques, where symbolic approaches can make contributions, e.g., for general intelligent agents. A separate Manifesto will provide an assessment of the area, and will give a set of recommendations regarding the future of KR and its promotion.

6.41 Machine Learning and Logical Reasoning: The New Frontier

Organizers: Sébastien Bardin, Vijay Ganesh, Somesh Jha, and Joao Marques-Silva

Seminar No. 22291

Date: July 17–22, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.7.80

Participants: Rajeev Alur, Sébastien Bardin, Chih-Hong Cheng, Jonathan Chung, Judith Clymo, Artur d'Avila Garcez, Alhussein Fawzi, Marc Fischer, Pascal Fontaine, Matt Fredrikson, Vijay Ganesh, Sebastian Junges, Chunxiao (Ian) Li, Ravi Mangal, Georg Martius, Kuldeep S. Meel, Grégoire Menguy, Matthew Mirman, Anselm Paulus, Markus N. Rabe, Joseph Scott, Xujie Si, Armando Tacchella, Hazem Torfah, Caterina Urban, Saranya Vijayakumar

This Dagstuhl Seminar is meant to be the first in a series, bringing together researchers from the two main pillars of AI, namely, logical reasoning and machine learning (ML), with a sharp focus on solver-based testing, analysis, and verification (TAV) methods aimed at improving the reliability and security of ML-based systems, and conversely, the use of ML heuristics in improving the power of solvers/provers. A third, albeit smaller focus is neuro-symbolic reasoning (NSR), that aims to combine the power of ML to learn deep correlations with the ability of solvers to perform logical inference as applied to many domains (including but not limited to math and logic).

While many previous Dagstuhl Seminars focus on sub-fields of this particular seminar (SAT, SMT, CP or machine learning), we focus here on the synergies and interplay between all them. Our goal is to deepen the understanding of the connections between learning and reasoning, and draw mutually beneficial research directions.

General context: Bringing ML and Logic Reasoning Closer

Since its very inception, Artificial Intelligence (AI) has largely been divided into two broad fields, namely, machine learning (ML) and logical reasoning, that have developed relatively independent of each other. Each of these sub-fields has had a deep and sustained impact on many topics in computer science and beyond, despite the limited interaction between them over the years. However, in recent years new problems and opportunities have come to fore that point towards combinations of ML and logical reasoning as the way forward [1] [34]. In this seminar, we wanted to explore combinations of ML and logical reasoning, under the following three specific themes:

Logic Reasoning for ML. Neural Networks (NN) today are ubiquitous and are being deployed as part of critical civilian and defense infrastructure, business processes, automotive software, and governmental decision-making systems. Unfortunately, despite their efficacy in solving many problems, NNs are brittle, unreliable, and pose significant security/privacy challenges [2]. The question of safety and security of NNs has therefore become a great concern to scientists, companies, and governments. In response to this problem, a nascent field of TAV methods for NNs is developing [3]. Key research directions in this context include logics aimed at symbolically representing NNs and their properties [4], novel solving methods [5], as well as solver-based TAV techniques specifically tailored for NNs [6]. A related set of questions focus on explainability and interpretability of NNs [7]. Finally, researchers are also exploring methods that combine logical reasoning within NN learning processes, with the aim of making them adversarially robust [8, 9]. The seminar aimed to bring together leading researchers in these topics, enabling cross-fertilization of ideas at a critical juncture in the development of the field.

ML for Logic Reasoning. In recent years, there has been considerable effort aimed at developing ML-based heuristics for logic reasoning engines such as SAT, SMT, and CP solvers. The premise of this line of research is that logic solvers are a combination of methods that implement proof rules and ML-based heuristics aimed at optimally selecting, sequencing, and initializing such proof rules [10]. This has led to new efficient solving algorithms that can solve real-world formulas with millions of variables and clauses in them. One of the many questions that was explored

[34] It goes without saying that it is infeasible to consider all possible combinations of ML and logical reasoning in this seminar. Hence, we focus primarily on problems inspired by testing, analysis, and verification (TAV) of ML and ML-based heuristics for logic solvers, with some forays into neuro-symbolic (a.k.a., neural-symbolic) AI.

in the seminar is how can we further deepen and strengthen this relation between ML and reasoning methods. Yet another line of research being explored is that of replacing rule-based solvers with NN-based logic reasoning (e.g., NeuroSAT [11]). Finally, methods are being developed to combine rule-based methods with reinforcement learning to automatically prove mathematical conjectures [12]. The seminar aimed to foster deeper interaction and collaboration among researchers who are pioneers in this intersection of ML-based methods and logical reasoning.

Neuro-symbolic Reasoning. The field of neuro-symbolic reasoning (NSR) aims to combine NNs with symbolic reasoning for the purposes of improving reasoning for many domains (including but not limited to pure math or logic). While at a high-level the field of NSR and logic solvers (with ML heuristics) may seem similar, they employ very different kinds of techniques and have differing goals [1]. For example, NSR researchers have developed methods for translating logical representations of knowledge into neural networks. Others have developed neuro-symbolic methods for concept-learning, and yet others have recently applied NSR to program synthesis. Can these concept-learning methods be adapted to the setting of logic solvers? Could it be that graph neural network (GNN) based representations of mathematical knowledge are easier to analyze? The seminar aimed to bring these two disparate communities closer together, that otherwise rarely interact with each other. In a nutshell, the aim of the seminar was to foster cross-fertilization of ideas between the logic reasoning, TAV, and ML communities.

In-depth Description of Focus Areas

Logic Reasoning for ML. As stated above, the reliability, safety, and security of NNs is a critical challenge for society at large. An example of a specific problem in this context is that of adversarial input generation methods against NNs. Many methods have been proposed to address this question, from randomized defense mechanisms to adversarial training to symbolic analysis of NNs via solvers, such as Reluplex [5] that are specifically designed to reason about NNs with ReLU units. Another line of work proposes verification of Binarized Neural Networks (BNNs) via SAT solvers [6]. These initial forays into reasoning for ML bring to fore new challenges, especially having to do with scalability of solvers for NN analysis. Traditional solver methods that scale well for typical software systems, do not seem to scale well for NNs. For example, it is known that solvers, such as Reluplex, are capable of analyzing NNs with only a few thousand nodes. The pressing question of this area of research then is *"How can we develop methods that enable solvers to scale to NNs with millions of nodes in them?"*

A related question has to do with appropriate logics to represent NNs and their properties. Recent work by Soutedeh and Thakur suggests that NNs can be represented symbolically as piecewise linear functions, even though they may use non-linear activation functions such as ReLU [4]. This suggests that there may be efficient solving methods capable of analyzing very large NNs. Yet another question in this setting is how do logic-based methods aimed at testing and verifying NNs compare against hybrid methods that do not require translation of NNs into logic. What are the tradeoffs in this setting?

Another interesting direction where logic reasoning can play a role is in explainability and interpretability of ML models. While both these questions have been long studied in AI and are closely related, they take particular importance in the context of NNs. We say a ML model is explainable, if there is discernable causal relationship between its input and output. Explanations for the behavior of NN, when presented in symbolic form, can be analyzed and debugged using solvers. Researchers have also developed solver-based xAI methods that aim to provide explanations for behavior of NNs [7]. By contrast, interpretable models are ones that have mathematical guarantees regarding their approximation or generalization errors. Solvers can play a role in this context as well via methods for generating counterfactuals (or adversarial examples) [1].

Strong points:

- *Adversarial attacks and defense mechanisms*
- *Neural network testing, analysis, and verification methods*
- *Piecewise linear symbolic representation of NNs*
- *Solvers for NNs*
- *Logic-guided machine learning*
- *Adversarial training*
- *Logic-based explainability and interpretability of NNs*

ML-based Heuristics for Logic Solvers. In recent years, ML-based methods have had a considerable impact on logic solvers. The key premise of this line of research is that logic solvers are a combination of proof systems and ML-based heuristics aimed at optimally selecting, sequencing, and initializing proof rules with the goal of constructing short proofs (if one exists) [10]. A dominant paradigm in this setting is modeling branching heuristics as RL methods to solve the multi-arm bandit (MAB) problem [10]. While this connection seems quite natural today and MAB-style methods have been shown to be empirically powerful, an important question remains as to why these heuristics are effective for industrial instances. A theoretical answer to this question can open up new connections between ML and logic solvers. Another direction of research that has been explored is solving SAT using NNs, *a la* NeuroSAT [11]. Finally, higher-order theorem provers have been developed recently at Google and elsewhere that combine RL with logic reasoning in order to automatically prove theorems from a variety of mathematical fields [12,13]. The seminar focused on these recent developments and the next steps in the research on combinations of ML-based methods with logic reasoning with the goal of achieving greater solver efficiency as well as expressive power.

Strong points:

- *ML-techniques for branching and restarts in SAT, SMT, and CP solvers*
- *Supervised learning methods for splitting and initialization in solvers*
- *NN-based methods for logical reasoning*
- *RL for higher-order theorem proving*

Neuro-symbolic Reasoning. Researchers in neuro-symbolic reasoning (NSR) have been independently developing algorithms that combine ML with symbolic reasoning methods with a slightly different focus than solver and theorem prover developers. NSR research has been focused on concept learning in a broader setting than math or logic, and the cross-fertilization of these ideas with logic-based methods can have deep impact both on NSR as well as solver research [1]. One of the key ideas we wanted to explore in this context is that of concept learning, i.e., learning of relations or concepts represented in a logical language directly from data. One interesting direction to explore would be how we can incorporate these methods in logic solvers? Another direction is to explore the synergy between NSR and synthesis of programs from examples. The seminar focused on bringing NSR and solver researchers closer together, given that they rarely interact in other settings.

Strong points:
- *Concept-learning, with possible applications in higher-order theorem provers*
- *Neuro-symbolic methods for program synthesis*
- *Concept learning for predicate abstraction*

Goals of the Seminar

The aim of this seminar was to bring together the logic reasoning and ML communities, thus shaping and setting the research agenda for ML-based solvers, TAV methods aimed at NNs, and NSR for many years to come.

The seminar highlighted the current challenges with symbolic analysis of NNs, scalability issues with solvers tailored for NNs, state-of-the-art ML-based heuristics for solvers, adapting NSR ideas to the setting of solvers and vice-versa, as well as bring to fore competing TAV methods that don't necessarily rely on symbolic representation of NNs.

Research questions. We highlight some of the main challenges at the intersection of ML and logic reasoning that were addressed during the seminar from different research perspectives, and discuss how we sought to combine or adapt current techniques to attack them.

- **Symbolic representation of NNs:** Recent work suggests that, while NNs are non-linear functions, they can be effectively modelled symbolically as piecewise linear functions. This is a significant advance since it dramatically simplifies the design of solvers for analyzing NNs. Some of the challenges that remain are algorithmic, i.e., how can NNs be efficiently converted into a symbolic representation.
- **Solvers for NNs:** As of this writing, Reluplex and its successors seem to be among the best solvers for analyzing the symbolic representations of NNs. Unfortunately, these tools scale to NNs with at most a few thousand nodes. There is an urgent need for novel ideas for solving algorithms that enable us to scale to real-world NNs with millions of nodes. Can hybrid methods that combine ML techniques with solvers scale more effectively than pure logic methods?
- **Combining Constraints and NN Learning:** Another set of questions we addressed is how can we improve the process via which NNs learn using logical constraints. In other words, can the back propagation algorithm be modified to take constraint or domain-specific knowledge into account? Can NNs be combined with logic solvers in a CEGAR-style feedback loop for the purposes of adversarial training?
- **Next steps in ML-based Heuristics for Solvers:** As stated earlier, modern solvers rely in significant ways on ML-based heuristics for their performance. We plan to focus on how we could strengthen this interaction further. For example, are there supervised learning methods for improving the performance of divide-and-conquer parallel SAT solvers. Can we develop ML-based methods for clause sharing in portfolio solvers? How about ML-based restarts and clause deletion policies?
- **Reinforcement learning (RL) and Theorem Provers:** There has been some recent success in combining basic RL methods with reasoning methods in the context of higher-order theorem provers. How can this combination be strengthened further to prove math theorems in a completely automated fashion?
- **Comparison of NN Verification with Testing and Fuzzing Methods:** Researchers have developed a variety of fuzzing methods aimed at NNs. These methods often scale better than verification techniques. On the other hand, unlike verification, testing techniques do not give any guarantees. What are the tradeoffs in this context of complete verification vs. scalability? Can we develop hybrid methods and light-weight verification techniques?
- **Concept Learning and Solvers:** Can we lift the ideas of concept learning from NSR to the setting of solvers, especially in the context of higher-order and combinatorial mathematics?

Synergies. We have also identified the following potential synergies between the ML, Solver, TAV, and NSR communities and expect strong interactions around these points:

- ML researchers in general (and RL in particular) can help refine the ML-based methods used by solver developers;
- Solver developers can propose constraint-based learning strategies for NNs (e.g., combining constraints with gradient-descent in the back propagation algorithm);
- Researchers who work in the space of TAV for NN can benefit greatly by better understanding the realistic security and safety concerns of the ML community;
- Solver developer can substantially benefit by better understanding concept learning from NSR researchers.

Expected results and impact on the research community. One of the core goals of the seminar is to bring together the many different research communities that work in the logic reasoning and ML fields, who unfortunately rarely talk to each other. We believe that the exchange of ideas between them – each with their own methods and perspectives – will help accelerate the future development of combinations of ML and logic reasoning. In terms of concrete outcomes, we believe the workshop is likely to lead to several collaboration projects, especially between members of different communities working on similar or related problems. Common benchmarks and regular meeting forums have also been discussed and we expect some progress there as well.

References

1 Amel, Kay R. From shallow to deep interactions between knowledge representation, reasoning and machine learning. 13th International Conference Scala Uncertainity Mgmt (SUM 2019), Compiègne, LNCS. 2019.

2 Goodfellow, Ian J and Shlens, Jonathon and Szegedy, Christian. Explaining and harnessing adversarial examples. arXiv preprint arXiv:1412.6572. 2014

3 Leofante, Francesco and Narodytska, Nina and Pulina, Luca and Tacchella, Armando. Automated verification of neural networks: Advances, challenges and perspectives. arXiv preprint arXiv:1805.09938. 2018

4 Sotoudeh, Matthew and Thakur, Aditya V. A symbolic neural network representation and its application to understanding, verifying, and patching networks. arXiv preprint arXiv:1908.06223. 2019

5 Katz, Guy and Barrett, Clark and Dill, David L and Julian, Kyle and Kochenderfer, Mykel J. Reluplex: An efficient SMT solver for verifying deep neural networks. In CAV 2017

6 Narodytska, Nina and Kasiviswanathan, Shiva and Ryzhyk, Leonid and Sagiv, Mooly and Walsh, Toby. Verifying properties of binarized deep neural networks. AAAI 2018

7 Samek, Wojciech and Wiegand, Thomas and Müller, Klaus-Robert. . Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. arXiv preprint arXiv:1708.08296. 2017

8 Fischer, Marc and Balunovic, Mislav and Drachsler-Cohen, Dana and Gehr, Timon and Zhang, Ce and Vechev, Martin. Dl2: Training and querying neural networks with logic. ICML 2019

9 LGML: Logic Guided Machine Learning. Scott, Joseph and Panju, Maysum and Ganesh, Vijay. AAAI 2020.

10 Jia Hui Liang, Hari Govind V. K., Pascal Poupart, Krzysztof Czarnecki and Vijay Ganesh. An Empirical Study of Branching Heuristics Through the Lens of Global Learning Rate. In SAT 2017.

11 Selsam, Daniel and Lamm, Matthew and Bünz, Benedikt and Liang, Percy and de Moura, Leonardo and Dill, David L. Learning a SAT solver from single-bit supervision. arXiv preprint arXiv:1802.03685. 2018

12 Kaliszyk, Cezary and Urban, Josef and Michalewski, Henryk and Olšák, Miroslav. Reinforcement learning of theorem proving. In Advances in Neural Information Processing Systems. 2018

13 Bansal, Kshitij and Loos, Sarah and Rabe, Markus and Szegedy, Christian and Wilcox, Stewart. HOList: An environment for machine learning of higher order logic theorem proving. ICML 2019.

6

6.42 Computational Approaches to Digitised Historical Newspapers

Organizers: Antoine Doucet, Marten Düring, Maud Ehrmann, and Clemens Neudecker

Seminar No. 22292

Date: July 17–22, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.7.112

Participants: Kaspar Beelen, Estelle Bunout, Sally Chambers, Simon Clematide, Mariona Coll-Ardanuy, Mickaël Coustaty, Marten Düring, Maud Ehrmann, Laura Hollink, Stefan Jänicke, Axel Jean-Caurant, Dario Kampkaspar, Jana Keck, Yves Maurer, Clemens Neudecker, Julia Noordegraaf, Eva Pfanzelter, David A. Smith, Martin Volk, Lars Wieneke

Context

For long held on library and archive shelving, historical newspapers are undergoing mass digitisation and millions of facsimiles, along with their machine-readable content captured via optical character recognition (OCR), are becoming accessible via a variety of online portals.[35] While this represents a major step forward in terms of preservation and access, it also opens up new opportunities and poses timely challenges for both computer scientists and humanities scholars [1–3, 14].

As a direct consequence, the last ten years have seen a significant increase of academic research on historical newspaper processing. In addition to decisive grassroots efforts led by libraries to improve OCR technology,[36] individual works dedicated to the development and application of tools to digitised newspaper collections have multiplied [10–13], as well as events such as evaluation campaigns or hackathons [5–8].[37] Besides, several large consortia projects proposing to apply computational methods to historical newspapers at scale have recently emerged, including ViralTexts[38], Oceanic Exchanges[39], *impresso* – Media Monitoring of the Past[40], NewsEye[41], Living with Machines[42], and Data-kbr-be[43] [9].

This momentum can be attributed not only to the long-standing interest of scholars in newspapers coupled with their recent digitisation, but also to the fact that these digital sources concentrate many challenges for computer science, which are all the more difficult – and interesting – since addressing them requires taking digital (humanities) scholarship needs and knowledge into account. Within interdisciplinary frameworks, various and complementary approaches spanning the areas of natural language processing, computer vision, large-scale computing and visualisation, are currently being developed, evaluated and deployed. Overall, these efforts are contributing a pioneering set of tools, system architectures, technical infrastructures and interfaces covering several aspects of historical newspaper processing and exploitation.

Objectives

The aim of the seminar was to bring together researchers and practitioners involved in computational approaches to historical newspapers to share experiences, analyse successes and shortcomings, deepen our understanding of the interplay between computational aspects and digital scholarship, and begin to design a road map for future challenges. Our seminar was guided by the vision of methodologically reflected, competitive and sustainable technical frameworks capable of providing an

[35] Such as those discussed in [16] and by this seminar's working group on transparency and fairness, see Section 4.3 in the full report.
[36] See e.g., the OCR-D project, an ecosystem for improving OCR on historical documents: https://ocr-d.de/en/about and [4].
[37] See the 2017 edition of the Coding Da Vinci cultural hackathon or the 2019 edition of the Helsinki Digital Humanities Hackathon.
[38] A project aiming at mapping networks of reprinting in 19th-century newspapers and magazines (US, 2012–2016): https://viraltexts.org
[39] A project tracing global information networks in historical newspaper repositories from 1840 to 1914 (US/EU, 2017–2019): https://oceanicexchanges.org
[40] A project which tackles the challenge of enabling critical text mining of newspaper archives (CH, 2017–2020): https://impresso-project.ch
[41] A digital investigator for historical newspapers (EU, 2018–2022): https://www.newseye.eu
[42] A project which aims at harnessing digitised newspaper archives (UK, 2018–2023): https://www.turing.ac.uk/research/research-projects/living-machines
[43] A project which aims at facilitating data-level access to KBR's collections (mainly newspapers) for open science (2020–2022):https://www.kbr.be/en/projects/data-kbr-be/.

extensive, sophisticated and possibly dynamic access to the content of digitised historic newspapers in a way that best serves the needs of digital scholarship. We are convinced that in order to meet the many challenges of newspaper processing and to accommodate the demands of humanities scholars, only a global and interdisciplinary approach that looks beyond technical solutionism and embraces the complexity of the source and its study can really move things forward.

Participants and Organisation

The seminar gathered 22 researchers[44] with backgrounds in natural language processing, computer vision, digital history and digital library, the vast majority of whom had previously worked on historical newspapers and were familiar with interdisciplinary environments. To structure and coordinate the work of the seminar, the organisers proposed a mixture of plenary sessions, working groups, and talks, as follows (see also Fig. 6.5):

- **Spotlight talks** on day 1, where each participant briefly introduced him/herself and gave an opinion or statement on his/her current view of the main topic of the seminar (3-minute/1-slide).
- **Demo session**, where some participants shortly introduced a relevant asset (e.g., a dataset, tool, interface, on-going experiment).
- **Working group sessions**, during which groups composed of computer scientists and humanities scholars focused on a specific question. Work within a group featured different moments, with: expert group discussion, where people with similar backgrounds exchanged in order to align their understanding of the question at hand and to prioritise problems; observation of concrete research and workflow practices on existing approaches and/or tools; cross-interviews, where people from one domain interviewed one person from another domain about a specific point; mixed group discussion, where everybody jointly reflected; and writing time, where the group wrote a report summarising its findings.
- **Reporting sessions**, where working groups reported their discussion and presented their main conclusions and recommendations in a plenary session.
- **Morning presentations**, where researchers shared their experience from a project and/or their view on a specific topic, followed by a discussion with the participants.
- **Evening talks**, where researchers shared their experience and views on a topic at large. We proposed three evening lectures that addressed the field of digitised newspapers from the perspective of computer science, digital history, and digital libraries.

<table>
<tr><th></th><th>Monday</th><th>Tuesday</th><th>Wednesday</th><th>Thursday</th><th>Friday</th></tr>
<tr><td>9:00 - 10:00</td><td rowspan="3">- General introduction
- Spotlight talks (ca. 22*3min) with coffee break
- Topics presentation and expressions of interest
- Identification of groups
(please refer to the Wiki for detailed schedule)</td><td>Morning Talks (2* 20+10min)</td><td rowspan="2">Demo Session</td><td>Group Reporting</td><td rowspan="2">Group Reporting & Working Group Session</td></tr>
<tr><td>10:00 - 11:00</td><td>Group Reporting (30')</td><td rowspan="2">Working Group Session</td></tr>
<tr><td>11:00 - 12:00</td><td>Working Group Session</td><td>Working Group Session</td><td>Concluding session</td></tr>
<tr><td>12:15 - 13:30</td><td colspan="4">lunch</td><td>end of seminar</td></tr>
<tr><td>13:30 - 15:30</td><td>Room assignment and working group session</td><td>Working group session</td><td rowspan="5">excursion</td><td>Working Group Session</td><td rowspan="5"></td></tr>
<tr><td>15:30 - 16:00</td><td>coffee break</td><td>coffee break</td><td>coffee break</td></tr>
<tr><td>16:00 - 17:30</td><td rowspan="2">Working Group Session</td><td rowspan="2">Working Group Session</td><td rowspan="2">Working Group Session</td></tr>
<tr><td>17:30 - 18:00</td></tr>
<tr><td>18:00 - 19:00</td><td>dinner</td><td>dinner</td><td>dinner</td></tr>
<tr><td>20:15 - 21:15</td><td>Evening talk (computer science perspective)</td><td>Evening talk (digital history perspective)</td><td>outside dinner</td><td>Evening talk (digital library perspective)</td></tr>
</table>

Fig. 6.5
Schedule of the seminar.

Topics

The topics and modus operandi of the seminar were not set in stone but discussed and validated with all participants during the first day. First, the organisers proposed three main topics (and several corresponding sub-questions) for the participants to discuss and reflect on during the seminar. On this basis, participants were then invited to express the specific themes, questions and issues they wished to work on, in a traditional post-it session. Finally, these propositions were examined and structured by the organisers, who defined four working groups.

[44] Some participants had to cancel at the last moment due to the pandemic; we thank them for their initial commitment and hope that there will be future opportunities.

Proposed topics

As a starting point, organisers proposed to consider three closely intertwined topics, which are detailed below to further illustrate the background knowledge of this seminar.

1. **Document Structure and Text Processing.** While recent work on the semantic enrichment of historical newspapers has opened new doors for their exploration and data-driven analysis from a methodological perspective (e.g., n-grams, culturomics), results up to now often confirmed common knowledge and were not always considered relevant by historians. The next natural and eagerly-awaited step consists in enriching newspaper contents and structure with semantic annotations which allow for the exploration of far more nuanced research questions. In this regard, several issues arise, among others:
   - **Q1.1 –** ***Complex structures and heterogeneity of contents.*** Newspapers are typically composed of a diverse mix of content including text, image/graphical elements, as well as tabular data and various other visual features. The proper segmentation of the page content into individual information pieces is key for enabling advanced research and analysis. This includes the modelling and detection of logical units on the document (or specifically, issue) level as, e.g., articles can span across multiple pages. Also of high relevance to researchers is the more advanced classification and semantic labelling of content units, separating categories such as information, opinion, stock market indices, obituaries, humour, etc. Despite a growing interest, a good understanding of these complex structures as well as methods and technologies for identifying, classifying and accessing diverse content types through appropriate data models and search interfaces are still lacking.
   - **Q1.2 –** ***Diachronic processing.*** Historical newspaper material poses severe challenges for computational analysis due to their heterogeneity and evolution over time. At language level, besides historical spelling variation which leads to major problems in text recognition and retrieval, sequential labelling tasks such as named entity recognition and disambiguation are problematic and often require time-specific resources and solutions. At document level, text classification or topic modelling need to pay attention to the necessary historical contextualisation of their category schemes or corpus time-spans in order to avoid anachronisms. Finally, at structure level, layout processing faces similar challenges and its application needs to adapt to changing sources.
2. **Visualisation, Exploitation and Digital Scholarship.** Historians and other user groups require tools for content discovery and management to reflect their iterative, exploratory research practices. The opportunities and challenges posed by mass digitised newspapers and other digitised sources require them to adjust their current workflows and to acquire new skills.
   - **Q2.1 –** ***Transparency and digital literacy.*** In the context of research, humanists' trust in computer systems is dependent on sufficient comprehension of the quality of the underlying data and the performance of the tools used to process it. One way to generate such trust is to create transparency, here understood as: information on the provenance and quality of digital sources; information which allows users to make informed decisions about the tools and data they use; and information which allows their peers to retrace their steps. Such transparency empowers users to use the system in a reflective way. But there is to-date no shared understanding of which information exactly is required to achieve transparency: technical confidence scores are themselves hard to interpret and do not translate easily into actionable information. Likewise, historians can not be expected to be aware of the consequences of all the algorithmic treatments to which their digitised sources have been exposed. Instead, the identification of the most relevant biases and their concrete consequences for users appears to be a more realistic approach. Once these are understood, counter-action can be taken.
   - **Q2.2 –** ***Iterative content discovery and analysis.*** In contrast to many other applications in computer science, the discovery of relevant content is of greater interest to historians than the detection of patterns in datasets following a priori hypotheses. Historical research is typically iterative: The study of documents yields new insights which determine future exploration strategies and allow scholars to reassess the value of the sources they have consulted. Semantically enriched content offers multiple ways to support this iterative exploration process. New tools for content discovery also require "generous" interfaces, i.e., interfaces which allow users to discover content rather than relying on narrow keyword search [15].
3. **System Architecture and Knowledge Representation.** The application of various natural language processing and computer vision components which transform noisy and unstructured material into structured data also poses challenges in terms of system architecture and knowledge representation. If those two well-studied fields already offer a strong base to build upon, many questions arise from newspaper source specificities and the digital humanities context.
   - **Q3.1 –** ***Managing provenance, complexity, ambivalence and vagueness.*** Lots of factual and non factual information is extracted from newspaper material and needs to be stored and interlinked. In this regard, two points require great attention. First, newspapers – like any other historical source – represent past realities which do not necessarily align with present-day realities: institutions and countries change names or merge, country borders move or become disputed and professions change or disappear. These temporal shifts, ambivalences and contradicting information cause historical data to be highly complex and sometimes disputed, and the representation of this complexity poses interesting challenges for computer science. Second, if processing steps, and possibly intermediary representations, of algorithms are recorded for the purpose of transparency, this meta-knowledge needs to be stored alongside the data.
   - **Q3.2 –** ***Dynamic processing.*** Historical newspaper processing outputs are useless if not used by scholars who wish to investigate research questions. If all methods and practices can not be transposed as they stand from analogue to digital, careful consideration must be given to how best to accommodate scholarship requirements in digital environments where primary sources are turned into data.

Selected Topics and Working Groups

The discussion around topics led to the definition of four working groups which the participants joined on the first day (on a voluntary basis) and in which they worked throughout the week. No guidelines were given and the groups were free to adapt the

direction of their work. Each group wrote a report summarising their activities and findings in Section 4 of the full report.

1. **Working Group on Information Extraction.** Initiated around the topic of information extraction, this group eventually settled on the specific topic of person entity mentions found in historical newspapers but not present in knowledge bases, a.k.a. "hidden people". The group defined a number of challenges and worked – in a productive hackathon style – on several experiments (see Section 4.1 of the full report).
2. **Working Group on Segmentation and Classification.** Members of this group quickly discarded the segmentation question to focus on classification only, considering classification scope and practices in relation to digitised newspapers (see Section 4.2 of the full report).
3. **Working Group on Transparency, Critics and Newspapers as Data.** This group (the largest) worked on a set of recommendations regarding the different aspects of transparency and fairness needed for the analysis of digitised and enriched historical newspaper collections (see Section 4.3 of the full report).
4. **Working Group on Infrastructure and Interoperability.** This group discussed the issue of consolidation, growth and sustainability of current and future achievements in digitisation, access, processing and exchange of historical newspapers (see Section 4.4 of the full report).

Spotlight Talks on the Main Challenges Ahead for Digitised Historical Newspapers

On the first morning of the seminar, the organisers asked participants to briefly present their views on some questions they had to consider in advance. These questions were:

- What are the main challenges we need to address in relation with historical newspapers?
- What is the most exciting opportunity you would like to explore during this seminar?
- If you were given €1 million to spend in the next 6 months on historical newspapers, what would you do?

As well as being a good ice-breaker and kick-off to the seminar, the series of responses to these questions documents what a community of researchers in July 2022 believe to be the next challenges for computational approaches to historical newspapers. In total, 21 researchers formulated no less than 67 statements as responses to the first question. In what follows, we provide a summary of the main ideas and suggestions which we have grouped in 8 themes that cover more or less the whole spectrum of activities around digitised newspapers. Apart from this grouping, no further reflection or refactoring has been done on these statements. While most of the answers are not a surprise to those familiar with the subject, they confirm existing needs, reflect on-going trends, and reveal new lines of research.

► **Document processing.** A first group of statements relates to optical character recognition and optical layout recognition (OLR), two critical processes when working with newspapers. These two document image refinement techniques are extremely difficult when applied to such sources (especially for collections digitised long ago), which explains why they are still high on the agenda despite all the efforts invested in recent years. The views expressed highlight and confirm several dimensions, namely: OCR and OLR quality needs to be improved, finer-grained segmentation and classification of news items is necessary, and processes should be more robust across time and collections. Intensive work is being carried out in these areas.

Verbatim statements:

- *How to make available digitised newspaper collections in high quality OCR+layout* (Clemens Neudecker);
- *High quality article segmentation and classification* (Maud Ehrmann, Mickaël Coustaty);
- *The (massive) segmentation bottleneck* (Antoine Doucet);
- *Robust layout recognition* (Matteo Romanello);
- *A level playing field: OCR and fine-grained content segmentation* (Marten Düring);
- *Improving article segmentation (e.g., wrt advertisements and classifieds)* (Mariona Coll-Ardanuy);
- *Layout recognition (e.g., article separation, recognition of headings and authors' names)* (Dario Kampkaspar);
- *OCR+, layout recognition, article segmentation and classification* (Eva Pfanzelter);
- *Better article segmentation, and a way to deal with heterogeneous qualities of segmentation in DL* (Axel Jean-Caurant);
- *Quality of OCR across periods, languages and original document qualities* (Yves Maurer);
- *Standardised approaches to segment historical newspaper pages* (Stefan Jänicke).

► **Text and image processing.** This group encompasses all types of content processing applied to OCR and OLR outputs in view of enriching newspaper contents with further information, usually in the form of semantic annotations and item classification. The main challenges that emerge are: robustness (i.e., approaches that perform well on challenging, noisy input), finer-grained information extraction, few-short learning (to compensate lack of training data), transferability (approaches that perform well work across settings), multilinguality, multimodality, entity linking, interlinking of collections, and transmedia approaches.

Verbatim statements:

- *Robust multilingual information extraction* (Maud Ehrmann);
- *Developing methods that are robust to OCR errors* (Mariona Coll-Ardanuy);
- *Words with meaning change over time* (Martin Volk);
- *Text summarisation and text classification (monolingual and across languages)* (Martin Volk);
- *How to automatically detect genres (in particular film reviews and film listings)* (Julia Noordegraaf);
- *Multilinguality* (Eva Pfanzelter);
- *Ease multilingual scholarship (qualitative and quantitative)* (Antoine Doucet);
- *Investigating the relation between (or intertwining of) image and text* (Kaspar Beelen);
- *Embedding newspaper content within the media landscape* (Kaspar Beelen);
- *Data mining in newspapers (e.g., biographies, TV/radio programmes)* (Marten Düring);
- *Robust entity linking (multilingual historical documents)* (Matteo Romanello);
- *Entity Linking and visualisations over time, space and networks* (Martin Volk);
- *Linking with other data sources (parliamentary protocols, wiki-data, (historic) names-db, other newspaper portals, (historic) place names db, etc.)* (Eva Pfanzelter);
- *Create links between newspaper contents (topics, entities) and knowledge bases* (Simon Clematide);
- *Automate content analysis (discourses, argumentation, events, meaning, topics) to enable historical research* (Eva Pfanzelter);

– *Learn with few samples and human interactions* (Mickaël Coustaty).

► **Digitisation and Content Mining Evaluation.** Here we have grouped together views on the evaluation of technical approaches and tools at large, and the means to implement it. Important points that emerge are: better metrics, more and diverse gold standards, and better contextualisation and understanding of (sources of) errors.

Verbatim statements:

– *How to arrive at common methods and metrics for quality of digitised newspapers* (Clemens Neudecker);
– *Sustainably sharing ground truth datasets and training models* (Sally Chambers);
– *Developing a variety of NLP benchmarks for different tasks across different languages and types of publications* (Mariona Coll-Ardanuy);
– *Build a general taxonomy of content items (including ads, service tables, etc.) and prepare well-sampled data sets from a variety of publication places and time periods* (Simon Clematide);
– *Disentangling correlation of errors and missingness with time, place, language, network position, etc.* (David Smith).

► **Exploration of (enriched) newspaper collections and beyond.** One of the opportunities that researchers have been working on in recent years is new ways of exploring newspaper content. This group of statements is part of this context and highlights some of the long-awaited next steps: unified access to newspaper collections, support for data-driven research, and connection to other archives.

Verbatim statements:

– *Access to newspaper content across collections/projects/platforms* (Matteo Romanello);
– *Unified access to all collections with advanced exploration capacities* (Maud Ehrmann);
– *A unified framework to REALLY make collections accessible, usable and interoperable* (Antoine Doucet);
– *Access across collections and copyright hurdles* (Marten Düring);
– *Silos* (Jana Keck);
– *Data-driven linking and analysis of multiple types of sources (e.g., radio, TV, parliamentary records) and datasets (e.g., land ownership, migration)* (Marten Düring);
– *User-driven (from novice to expert) image, information and metadata etc. extraction* (Eva Pfanzelter);
– *Offer our users more than search, but what? Topics, recommenders, . . . ?* (Yves Maurer);
– *Contextual information extracted from the corpus: hints on rubrics, themes, top keyword per month etc.* (Estelle Bunout);
– *Comparative analyses of contents (political targets of publishers), ordering of articles and time-based development of topics* (Stefan Jänicke).

► **Working with data.** In addition to working with enriched sources that can be semantically indexed and thus more easily retrieved and analysed, researchers (especially historians) also express the wish to work directly with raw data – digitised documents, annotations, or both – and be able to build their own corpora.

Verbatim statements:

– *How to create useful datasets and corpora from digitised newspapers* (Clemens Neudecker);
– *Availability of digitisation output (images, text) for further use (beyond interfaces)* (Estelle Bunout);
– *Newspapers as Data: how to facilitate dataset / corpus building* (Sally Chambers);
– *Ease the building and sharing of corpora, taking into account the context of creation (queries, quality, etc)* (Axel Jean-Caurant).

► **Collections, source and tool criticism; Documentation; Inclusivity.** The validity of any conclusions drawn in empirical research depends on a solid understanding of the data used for the analysis. Digitised and enriched newspapers contain multiple levels of processing which often vary significantly across titles in terms of processing quality and extent of enrichment. The statements below point to key challenges and opportunities to advance our reflected analysis of digitised newspapers.

Verbatim statements:

– *How to support and perform source criticism on digitised newspaper collections* (Julia Noordegraaf);
– *Methodological guidelines for the computational analysis of newspaper content* (Julia Noordegraaf);
– *Describing how biases arise in digitised newspapers collections ("full-stack bias")* (Kaspar Beelen);
– *Understanding how structured missingness and data quality affect (historical) research* (Kaspar Beelen);
– *Selection criteria guidelines for what is being selected, digitised, accessible and how it is represented, searchable, and available* (Jana Keck);
– *Trustable and/or understandable approaches to meet users' needs* (Mickaël Coustaty);
– *How do we make collections as well as access mechanisms inclusive?* (Laura Hollink);
– *How do we monitor fairness of computational approaches to historic newspapers?* (Laura Hollink);
– *How well does the collection support different user groups?* (Laura Hollink);
– *How do we make perspectives in the data explicit? (e.g., in NL context: words signalling a colonial perspective)* (Laura Hollink);
– *Information on the scope, contents and quality of a collection, e.g., included titles, covered time periods, granularity of items (page vs. article), OCR quality, corpus statistics* (Estelle Bunout);
– *Investigate the role of attributes like font face and style, margins, layout, paper, etc.* (Stefan Jänicke);
– *Book-historical studies of editorship and publishing (costs, layout, format, advertising, syndicates, networks) crossing national and cataloguing (newspaper/magazine) boundaries* (David Smith);
– *Investigate the role of attributes like font face and style, margins, layout, paper, etc.* (Stefan Jänicke).

► **Workflows.** The combination of multiple processes, moreover between different actors, requires the design of more standardised and efficient workflows encompassing the many processing steps that have emerged in recent years.

Verbatim statements:

– *Advanced digitisation workflows: from digitisation to OCR to article segmentation* (Sally Chambers);
– *Workflows that conflate search, annotation, classification, corpus construction* (David Smith).

► **Legal matters.** Finally, a last set of (unavoidable) challenges concerns legal issues, with questions of copyright clearance and management, and of personal data, whether it be

user data handled by platforms, or the right to be forgotten.

Verbatim statements:

- *Find sustainable ways to work with copyright-restricted data sets* (Yves Maurer);
- *Access across collections and copyright hurdles* (Marten Düring);
- *Copyrights and proprietary rights, image rights etc.* (Eva Pfanzelter);
- *Legal questions (copyright, personal rights, etc.)* (Dario Kampkaspar).

Acknowledgements

This seminar was originally planned for September 2020 but was cancelled due to the COVID-19 pandemic and rescheduled 2 years later. We would like to thank the administrative and scientific teams at Dagstuhl for their support and professionalism throughout the (re)organisation of this seminar as well as the staff on site for their valuable every day help and care. We also thank all the participants for accepting our invitation to spend a week exchanging views, examining, questioning, debating (and writing) about computational approaches to historical newspapers.

References

1 Bingham, A. The Digitization of Newspaper Archives: Opportunities and Challenges for Historians. *Twentieth Century British History*.21, 225–231 (2010,6)

2 Deacon, D. Yesterday's Papers and Today's Technology: Digital Newspaper Archives and "Push Button" Content Analysis. *European Journal Of Communication*. 22, 5–25 (2007)

3 Nicholson, B. The Digital Turn. *Media History*. 19, 59–73 (2013,2)

4 Neudecker, C., Baierer, K., Federbusch, M., Boenig, M., Würzner, K., Hartmann, V. & Herrmann, E. OCR-D: An End-to-End Open Source OCR Framework for Historical Printed Documents. *Proceedings Of The 3rd International Conference On Digital Access To Textual Cultural Heritage*. pp. 53–58 (2019,5)

5 Ehrmann, M., Romanello, M., Najem-Meyer, S., Doucet, A. & Clematide, S. Extended Overview of HIPE-2022: Named Entity Recognition and Linking in Multilingual Historical Documents. *Proceedings Of The Working Notes Of CLEF 2022 – Conference And Labs Of The Evaluation Forum*. 3180 (2022) https://infoscience.epfl.ch/record/295816

6 Ehrmann, M., Romanello, M., Flückiger, A. & Clematide, S. Extended Overview of CLEF HIPE 2020: Named Entity Processing on Historical Newspapers. *Working Notes Of CLEF 2020 – Conference And Labs Of The Evaluation Forum*. 2696 pp. 38 (2020)

7 Clausner, C., Antonacopoulos, A., Pletschacher, S., Wilms, L. & Claeyssens, S. PRImA, DMAS2019, Competition on Digitised Magazine Article Segmentation (ICDAR 2019). (2019), https://www.primaresearch.org/DMAS2019/

8 Rigaud, C., Doucet, A., Coustaty, M. & Moreux, J. ICDAR 2019 Competition on Post-OCR Text Correction. *2019 International Conference On Document Analysis And Recognition (ICDAR)*. pp. 1588–1593 (2019)

9 Ridge, M., Colavizza, G., Brake, L., Ehrmann, M., Moreux, J. & Prescott, A. The Past, Present and Future of Digital Scholarship with Newspaper Collections. *DH 2019 Book Of Abstracts*. pp. 1–9 (2019), http://infoscience.epfl.ch/record/271329

10 Kestemont, M., Karsdorp, F. & Düring, M. Mining the Twentieth Century's History from the Time Magazine Corpus. *Proceedings Of The 8th Workshop On Language Technology For Cultural Heritage, Social Sciences, And Humanities (LaTeCH)*. pp. 62–70 (2014)

11 Lansdall-Welfare, T., Sudhahar, S., Thompson, J., Lewis, J., Team, F. & Cristianini, N. Content Analysis of 150 Years of British Periodicals. *Proceedings Of The National Academy Of Sciences*. **114**, E457-E465 (2017)

12 Moreux, J. Innovative Approaches of Historical Newspapers: Data Mining, Data Visualization, Semantic Enrichment. *IFLA News Media Section, Lexington, August 2016, At Lexington, USA*. (2016,8), https://hal-bnf.archives-ouvertes.fr/hal-01389455

13 Wevers, M. Using Word Embeddings to Examine Gender Bias in Dutch Newspapers, 1950–1990. *Proc. Of The 1st International Workshop On Computational Approaches To Historical Language Change*. (2019), https://www.aclweb.org/anthology/W19-4712

14 Bunout, E., Ehrmann, M. & Clavert, F. (editors) Digitised Newspapers – A New Eldorado for Historians? Tools, Methodology, Epistemology, and the Changing Practices of Writing History in the Context of Historical Newspaper Mass Digitisation. De Gruyter (2022, in press), doi:10.1515/9783110729214, https://www.degruyter.com/document/isbn/9783110729214/html

15 Whitelaw, M. Generous Interfaces for Digital Cultural Collections. *Digital Humanities Quarterly*. **9** (2015), http://www.digitalhumanities.org/dhq/vol/9/1/000205/000205.html

16 Ehrmann, M., Bunout, E. & Düring, M. Historical Newspaper User Interfaces: A Review. *Proceedings Of The 85th International Federation Of Library Associations And Institutions (IFLA) General Conference And Assembly*. pp. 24 (2019), https://infoscience.epfl.ch/record/270246?ln=en

6.43 Algorithmic Aspects of Information Theory

Organizers: Phokion G. Kolaitis, Andrej E. Romashchenko, Milan Studený, and Dan Suciu

Seminar No. 22301

Date: July 24–29, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.7.180

Participants: Marcelo Arenas, Albert Atserias, Amos Beimel, Tobias Andreas Boege, Janneke Bolt, Laszlo Csirmaz, Kyle Deeds, Oriol Farras, Yuval Filmus, Emirhan Gürpinar, Miika Hannula, Peter Harremoës, Batya Kenig, Phokion G. Kolaitis, Rostislav Matveev, Fabio Mogavero, Carles Padró, Andrei Romashchenko, Sudeepa Roy, Alexander Shen, Milan Studený, Dan Suciu, John MacLaren Walsh, Lele Wang, Geva Yashfe

The goal of this seminar was to bring together researchers from several communities who share an interest in the methods and the uses of information theory. Participants included experts in information theory, databases, secret sharing, algorithms, and combinatorics. There were four tutorials, two from the information theory community and two from the database community, that helped define a common language and a common set of problems. There were several contributed talks, from experts in all these fields. The proof of one of the major open problems in information theory was announced at the workshop by not one, but, by *two* researchers, namely Cheuk Ting Li and Geva Yashfe, who used quite different techniques to independently solve this open problem. Overall, the workshop was a success.

Organization of the Seminar

The seminar was held between July 25–29, 2022 (Monday to Friday), and had 25 on-site participants, and 8 remote participants. Since the participants represented quite diverse communities, we started the first day with an introduction of each participant. The four tutorials were scheduled during the first two days: two tutorials on information inequalities and conditional independence were given by László Csirmaz and Milan Studený, and two tutorials on different aspects of database theory were given by Marcelo Arenas and Hung Ngo. All four tutorials were very well received, with many questions and lively discussions during and after the tutorials. There were 18 contributed talks in total, spread over all 5 days of the seminar. We scheduled two sessions to discuss open problems: one on Tuesday afternoon, and one on Thursday afternoon. The seminar concluded with an hour-long discussion assessing the seminar and contemplating future directions. Our collector, Tobias Boege, recorded all open problems, and later typed them for inclusion in the full report.

Outcomes of the Seminar

There are several major outcomes:

- Having participants with very diverse backgrounds enabled us to exchange interesting ideas and problems. Information theorists became inspired by problems that arise in database research, while database theoreticians learned tools and techniques from information theory; almost all talks raised algorithmic questions that inspired people from the algorithms community.
- We have assembled a list of open problems, which we included in the full report, and we also plan to publish independently. We hope that this list will help define the community interested in the algorithmic aspects of information theory, and will also inspire young researchers to contribute to this emerging area.
- At the end of the workshop the participants expressed a lot of interest in continuing to have some organized forum for discussing problems in information theory. One of us (Andrei Romashchenko) is planning to organize regular talks, to be made publicly available online, via Zoom.
- Everyone was happily surprised that one major problem in information theory was essentially settled during this seminar. The problem asks whether the implication problem for conditional independence statements is decidable. This problem has been studied since at least the early 80's, and has resisted any prior attempts at settling it. Cheuk Ting Li announced a proof of the undecidability of this problem, and presented the high-level structure of the proof; he had

posted on arXiv a paper describing the proof just a few weeks before the seminar. Geva Yashfe had solved a different open problem: he showed that it is undecidable whether a given $(2^n - 1)$-dimensional vector is an almost entropic vector. Through discussions at the seminar, he realized that his proof can be extended to also prove that the implication problem for conditional independences is undecidable. He gave a presentation of his proof on the blackboard, during the seminar.

Acknowledgements

We are grateful to the Scientific Directorate and to the staff of the Schloss Dagstuhl – Leibniz Center for Informatics for their support of this seminar. We also wish to express our sincere thanks to Dr. Tobias Boege for collecting the abstracts of the talks and compiling the list of open problems.

6.44 Educational Programming Languages and Systems

Organizers: Neil Brown, Mark J. Guzdial, Shriram Krishnamurthi, and Jens Mönig

Seminar No. 22302

Date: July 24–29, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.07.205

Participants: Kenichi Asai, Michael Ball, Neil Brown, Youyou Cong, Barbara Ericson, Kathi Fisler, Diana Franklin, Elena Leah Glassman, Mark J. Guzdial, Bastiaan Heeren, Felienne Hermans, Jadga Hügle, Johan Jeuring, Amy Ko, Tobias Kohn, Shriram Krishnamurthi, Michael J. Lee, Eva Marinus, Jens Mönig, R. Benjamin Shapiro, Janet Siegmund, Ethel Tshukudu

To the world at large, programming is one of the most visible and valuable aspects of computing. Yet learning to program has been well documented as challenging and a barrier to entry. The way that a computer interprets a program literally and the complex interdependencies within programs are surprising to novices, who have a rather limited understanding of things that can go wrong in programs and what can be done to protect against these problems.

Thinking about computing (and specifically programming) education is particularly timely for several reasons:

- Numerous countries, states, and other geographic entities are making a big push to put computing into curricula for all students.
- Computing tools are now making a serious dent into several disciplines, not just traditional sciences (like physics and biology) but also in social sciences (like history and sociology). New hybrid areas like bioinformatics and data science are being created. People in numerous disciplines now would benefit from, and in some cases need to, learn how to program.
- People outside traditional academic structures are being given the power to program. Everything from spreadsheets to home automation systems are providing "scripting" interfaces that people can use to simplify or enrich their lives. At the same time, those who do not adapt to these trends risk being left behind in their jobs.

In response to the challenges that programmers face, computer scientists have created numerous programming languages and systems, such as interactive development environments (IDEs). To a professional, the more the tools, usually, the better. Tools that can compute advanced analyses, whose meaning may take a great deal of training to understand (as just one example, a worst-case execution time analysis to aid in building a real-time system, or a dependent type system that can statically analyze rich program invariants) can be well worth the investment.

Beginners, however, have a very different set of concerns. For instance:

Syntax Basic matters of syntax can be problematic. Beginners can struggle with the notion that the computer requires very precise utterances (something they have not had to deal with before in writing, math, and other disciplines, where a human reader is usually able to deal with ambiguity, and is forgiving). They may also struggle with basic tasks like typing.

Graduated Introduction A typical textbook will introduce concepts slowly and gradually: each chapter, for instance, might introduce one new concept. Our tools, however, do not typically offer this same graduation. Instead, a student put in front of (say) a Java IDE must confront all of Java: a typo could result in an indecipherable error, strange behavior, and so on. This can result in a very confusing, and unfriendly, learning environment.

Errors Computing is relatively unique in confronting learners with a large number of errors. But errors are intimidating to beginners (many of whom worry that they can "break the computer"), and can at any rate be a deflating experience. At the same time, errors can be viewed as a learning opportunity. How do we design errors – and more broadly, system feedback – in such a way that it is constructive, comprehensible, and encouraging?

Accessibility Computing has traditionally had a rather shameful relationship with users who have special needs. For example, the number of blind professional developers is negligible, and is not at all representative of the percentage of blind people in the population. However, spending a few minutes with our programming tools will make clear why this is not surprising at all: so much is oriented towards visual inspection and manipulation. Similarly, our tools are rarely tested against, for instance, the needs and capabilities of learners with cognitive impairments.

While it may take a long time before the entire computing "pipeline" can adjust to such needs, we can still start to make progress in this direction. Furthermore, the community of beginners who do not need much computing sophistication can immediately benefit from advances in this area.

Our seminar was successful in bringing together attendees with expertise in computing education, programming language design, and human-computer interaction. The presentation and discussions explored a wide range of issues, from language and environment design, to teaching methods and assessment issues. In particular, we studied and discussed:

Tools and Languages A wide variety of tools and languages were presented – from block-based to textual, from imperative to functional, to those embedded in games, and to those explicitly designed for programmers whose native language is not English. Some of the tools were explicitly for learning and teaching of programming, like tutors and ebooks. We discussed tools for understanding program execution, such as debuggers.

Blocks Many of the attendees are exploring block-oriented programming in some way, so we had several sessions focused on novel uses of block-based programming and applying the blocks modality in different kinds of programming languages and paradigms.

Data Increasingly, we recognize that learning programming is not just about the language and the IDE, but also about *data* – what data students use and how data are described and structured. Data can be motivating for students. Carefully selected data sets can play an important role in supporting learning in contexts outside of computing, e.g., about social issues or about scientific phenomena). Data can be complex and messy, which can take more time to explain and "clean."

Learning Issues Our discussion included the cognitive issues when learning programming and the challenges in helping students to transfer knowledge to new contexts. We discussed the strategies to be taught to students to help them succeed at reading, writing, and debugging programs. We saw teaching techniques that are unique to programming, like Parsons problems. We particularly focused on the cognitive tasks of planning and identifying goals, which are critical to student success in programming. We discussed learning trajectories that explain how we might expect student learning to occur.

Process Computer science has been called the study of algorithmic processes, and computer science education research needs to also be concerned with students learning computational processes in the context of other processes. Some of the processes we discussed include learning processes (i.e., what has to happen to make sure that learning is successful and is retained and possibly transferred), classroom processes (i.e., how does programming fit into the classroom context, including how work is evaluated), and student programming processes (i.e., students are learning design, development, and debugging processes, and need scaffolding to help them succeed and develop their processes to be more expert-like).

Practical Issues of Sustainability Building software has always been difficult to fit into most national research infrastructures. Funding agencies are often reticent to pay for software development, let alone maintenance. Maintaining software over time is expensive but is critical to test educational hypotheses in ecologically-valid contexts (i.e., classrooms vs laboratories) and to gain the benefits of novel software implementations in education. A key area of maintenance of educational software is the fit between the curriculum, the teacher, the school context, and the software. Software may need to change as teaching goals and teachers change, which is a new and complex area for software maintenance.

6

6.45 Visualization and Decision Making Design Under Uncertainty

Organizers: Nadia Boukhelifa, Christopher R. Johnson, and Kristi Potter

Seminar No. 22331

Date: August 15–19, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.8.1

Participants: Michael Böttinger, Nadia Boukhelifa, Remco Chang, Oliver Deussen, Christina Gillmann, Amy Gilmer, Michael Gleicher, Hans-Christian Hege, Christoph Heinzl, Miriah Meyer, Kristi Potter, Maria Riveiro, Gerik Scheuermann, Johanna Schmidt, Thomas Schultz, Regina Maria Veronika Schuster, Jarke J. van Wijk, Tatiana von Landesberger, Stefan Hagen Weber

Remote Participants: Tushar Athawale, Mehdi Chakhchoukh, Michael Correll, Stephanie Deitrick, Jake Hofman, Amit Jena, Christopher R. Johnson, Matthew Kay, Robert Lempert, Paul Rosen, Han-Wei Shen, Brian Summa, Kathleen Warrell, Thomas Wischgoll, Xiaoru Yuan, Liang Zhou

Uncertainty is an important aspect to data understanding. Without awareness of the variability, error, or reliability of a data set, the ability to make decisions on that data is limited. However, practices around uncertainty visualization remain domain-specific, rooted in convention, and in many instances, absent entirely. Part of the reason for this may be a lack of established guidelines for navigating difficult choices of when uncertainty should be added, how to visualize uncertainty, and how to evaluate its effectiveness. Unsurprisingly, the inclusion of uncertainty into visualizations is a major challenge to visualization [1]. As work concerned with uncertainty visualization grows, it has become clear that simple visual additions of uncertainty information to traditional visualization methods do not appropriately convey the meaning of the uncertainty, pose many perceptual challenges, and, in the worst case, can lead a viewer to a completely wrong understanding of the data.

The goal of this Dagstuhl Seminar was to bring together experts with diverse knowledge of uncertainty visualization and comprehension toward building a foundation of accessible, practical knowledge that practitioners and researchers alike can rely on in addressing challenges related to uncertainty. Specifically, this seminar brought together leaders in the field of uncertainty visualization and communication, along with experts on quantification and practitioners and domain experts dealing with uncertainty on a daily basis. Drawing on the knowledge of the participants, the seminar worked toward goals of synthesizing disparate findings and approaches from across computer science and related literature, noting current practices surrounding uncertainty, and identifying unsolved problems in common workflows, and areas needing further study.

As a major result from the seminar, the following challenges and research topics in visualization and decision making under uncertainty have been identified:

- Applications,
- Human-centered uncertainty visualization (including how to support “feeling uncertain”),
- A design process for uncertainty visualization,
- Defining terms related to uncertainty,
- Algorithms and uncertainty quantification,
- Software dissemination,
- User studies,
- Ethics of uncertainty (when to include uncertainty information),
- Surveys of uncertainty-aware visual analytics, and
- Teaching uncertainty visualization.

The top three challenges were discussed in depth during this intensive three and a half days Dagstuhl Seminar as part of the break-out groups, and are further discussed in the full report. In particular, the break-out groups examined uncertainty visualization research challenges from three complemen-

tary perspectives: from an application viewpoint looking at how uncertainty visualization and assessment are used in many domains; from a human-centered perspective considering the needs and information of the viewer; and from a more theoretical stand focusing on the problem space for designing uncertainty visualization.

The seminar ended with a presentation from each group and discussions on the next steps. Interesting research questions and potential solutions were identified during the discussions, and plans were made to continue the collaboration. Details of the individual talks and break-out group discussions are provided in the full report.

References

1 Chris R. Johnson and Allen R. Sanderson. *A next step: Visualizing errors and uncertainty*. IEEE Computer Graphics and Applications, 2003, vol. 23, no 5, p. 6–10.

6.46 Differential Equations and Continuous-Time Deep Learning

Organizers: David Duvenaud, Markus Heinonen, Michael Tiemann, and Max Welling

Seminar No. 22332

Date: August 15–19, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.8.20

Participants: Hananeh Aliee, Jesse Bettencourt, Olivier Bournez, Joachim M. Buhmann, Johanne Cohen, Biswadip Dey, Remco Duits, David Duvenaud, Maurizio Filippone, James Foster, Colby Fronk, Jan Hasenauer, Markus Heinonen, Patrick Kidger, Diederik P. Kingma, Linda Petzold, Jack Richter-Powell, Lars Ruthotto, Jacob Seidman, Arno Solin, Sho Sonoda, Nils Thuerey, Michael Tiemann, Filip Tronarp, Yves van Gennip, Max Welling, Verena Wolf, Daniel Worrall

Deep models have revolutionised machine learning due to their remarkable ability to iteratively construct more and more refined representations of data over the layers. Perhaps unsurprisingly, very deep learning architectures have recently been shown to converge to differential equation models, which are ubiquitous in sciences, but so far overlooked in machine learning. This striking connection opens new avenues of theory and practice of continuous-time machine learning inspired by physical sciences. Simultaneously, neural networks have started to emerge as powerful alternatives to cumbersome mechanistic dynamical systems. Finally, deep learning models in conjecture with stochastic gradient optimisation has been used to numerically solve high-dimensional partial differential equations. Thus, we have entered a new era of continuous-time modelling in machine learning.

This change in perspective is currently gaining interest rapidly across domains and provides an excellent and topical opportunity to bring together experts in dynamical systems, computational science, machine learning and the relevant scientific domains to lay solid foundations of these efforts. On the other hand, as the scientific communities, events and outlets are significantly disjoint, it is key to organize an interdisciplinary event and establish novel communication channels to ensure the distribution of relevant knowledge.

Over the course of this Dagstuhl Seminar, we established strong contacts, communication and collaboration of the different research communities. We had an exchange of each community's best practices, known pitfalls and tricks of the trade. We tried to identify the most important open questions and avenues forward to foster interdisciplinary research. To this end, this seminar featured not only individual contributed talks, but also general discussions and "collaboration bazaars", for which participants had the possibility to pitch ideas for break-out project sessions to each other. In the break-out sessions, participants could discuss open problems, joint research obstacles, or community building work.

6.47 Power and Energy-Aware Computing on Heterogeneous Systems (PEACHES)

Organizers: Julian De Hoog, Kerstin I. Eder, Timo Hönig, Daniel Mosse, and Max Plauth

Seminar No. 22341

Date: August 21–26, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.8.31

Participants: Ahmed Ali-Eldin Hassan, Antonio Carlos Schneider Beck Filho, Ruzanna Chitchyan, Liliana Cucu-Grosjean, Julian De Hoog, Kerstin I. Eder, Christian Eichler, Michael Engel, Mathias Gottschlag, Daniel Gruss, Benedict Herzog, Timo Hönig, Henry Hoffmann, Henriette Hofmeier, Romain Jacob, Alex Jones, Fiodar Kazhamiaka, Maja Hanne Kirkeby, Sven Köhler, Julia Lawall, Laércio Lima Pilla, Tulika Mitra, Daniel Mosse, Frank Mueller, Simon Peter, Max Plauth, George Porter, Andreas Schmidt, Gunnar Schomaker, Wolfgang Schröder-Preikschat, Sibylle Schupp, Jennifer Switzer, Devesh Tiwari, Lukas Wenzel, Samuel Xavier-de-Souza

More than ever, emissions, carbon footprint, and other related environmental concerns are at the forefront of society, from several different perspectives. There is an urgent need to understand how **computing** fits into the broader picture of our planet's energy consumption and what is the role of computing in reducing our carbon footprint worldwide. This requires new ways of thinking across different domains, and necessitates highly energy-efficient hardware and software designs that adapt to changing operating conditions to become more efficient. Collaboration is increasingly required across the entire system stack – from system designers to programmers and operators.

The Dagstuhl Seminar 22341 on "Power and Energy-Aware Computing on Heterogeneous Systems (PEACHES)" brought together experts from computer science and computer engineering that share a common vision towards reducing carbon emissions both using innovative designs for computing systems and techniques that bridge the gap between hardware and software, as well as using computing systems to manage other environment-influencing systems. Five principal topic areas were discussed in working groups during the meeting: Energy transparency from hardware to software, Energy optimisation and management, Sustainability in computing, "Green Computing" hackathons, and Disruptive paradigms.

The full report documents the program and the outcomes of PEACHES.

6

6.48 Privacy in Speech and Language Technology

Organizers: Simone Fischer-Hübner, Dietrich Klakow, Peggy Valcke, and Emmanuel Vincent

Seminar No. 22342

Date: August 21–26, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.8.60

Participants: Lydia Belkadi, Zinaida Benenson, Martine De Cock, Abdullah Elbi, Zekeriya Erkin, Natasha Fernandes, Simone Fischer-Hübner, Ivan Habernal, Meiko Jensen, Els Kindt, Dietrich Klakow, Katherine Lee, Anna Leschanowsky, Pierre Lison, Christina Lohr, Emily Mower Provost, Andreas Nautsch, Olga Ohrimenko, Jo Pierson, Laurens Sion, David Stevens, Francisco Teixeira, Natalia Tomashenko, Marc Tommasi, Peggy Valcke, Emmanuel Vincent, Shomir Wilson

In the last few years, voice assistants have become the preferred means of interacting with smart devices and services. Chatbots and related language technologies such as machine translation or typing prediction are also widely used. These technologies often rely on cloud-based machine learning systems trained on speech or text data collected from the users. The recording, storage and processing of users' speech or text data raises severe privacy threats. This data contains a wealth of personal information about, e.g., the personality, ethnicity and health state of the user, that may be (mis)used for targeted processing or advertisement. It also includes information about the user identity which could be exploited by an attacker to impersonate him/her. News articles exposing these threats to the general public have made national headlines.

A new generation of privacy-preserving speech and language technologies is needed that ensures user privacy while still providing users with the same benefits and companies with the training data needed to develop these technologies. Recent regulations such as the European General Data Protection Regulation (GDPR), which promotes the principle of privacy-by-design, have further fueled interest. Yet, efforts in this direction have suffered from the lack of collaboration across research communities. This Dagstuhl Seminar was the first event to bring 6 relevant disciplines and communities together: Speech Processing, Natural Language Processing, Privacy Enhancing Technologies, Machine Learning, Human Factors, and Law.

After 6 tutorials given from the perspective of each of these 6 disciplines, the attendees gathered into cross-disciplinary working groups on 6 topics. The first group analyzed the privacy threats and the level of user control for a few case studies. The second group focused on anonymization of unstructured speech data and discussed the legal validity of the success measures developed in the speech processing literature. The third group devoted special interest to vulnerable groups of users in regard to the current laws in various countries. The fifth group tackled the design of privacy attacks against speech and text data. Finally, the sixth group explored the legal interpretation of emerging privacy enhancing technologies.

The reports of these 6 working groups, which are gathered in the full report, constitute the major result from the seminar. We consider them as a first step towards a full-fledged interdisciplinary roadmap for the development of private-by-design speech and language technologies addressing societal and industrial needs.

6.49 Interactive Visualization for Fostering Trust in ML

Organizers: Polo Chau, Alex Endert, Daniel A. Keim, and Daniela Oelke

Seminar No. 22351

Date: August 28–September 2, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.8.103

Participants: Gennady Andrienko, Natalia V. Andrienko, Emma Beauxis-Aussalet, Michael Behrisch, Rita Borgo, Simone Braun, Peer-Timo Bremer, Polo Chau, David S. Ebert, Mennatallah El-Assady, Alex Endert, Brian D. Fisher, Barbara Hammer, Daniel A. Keim, Steffen Koch, Jörn Kohlhammer, Rafael M. Martins, Laura Matzen, Daniela Oelke, Jaakko Peltonen, Adam Perer, Maria Riveiro, Tobias Schreck, Harald Schupp, Hendrik Strobelt, Alexandru C. Telea, Stef Van den Elzen, Michel Verleysen, Emily Wall

Artificial intelligence (AI), and in particular machine learning (ML) algorithms, are of increasing importance in many application areas. However, interpretability, understandability, responsibility, accountability, and fairness of the algorithms' results – all crucial for increasing humans' trust into the systems – are still largely missing. All major industrial players, including Google, Microsoft, and Apple, have become aware of this gap and recently published some form of Guidelines for the Use of AI. While it is clear that the level of trust in AI systems does not only depend on technical but many other factors, including sociological and psychological factors, interactive visualization is one of the technologies that has strong potential to increase trust into AI systems. In our Dagstuhl Seminar, we discussed the requirements for trustworthy AI systems including sociological and psychological aspects as well as the technological possibilities provided by interactive visualizations to increase human trust in AI. As a first step, we identified the factors influencing the organizational and sociological as well as psychological aspects of AI and partitioned them into relationship-based and evidence-based aspects. Next, we collected measures that may be used to approximate these aspects, such as interaction logs, eye tracking, and EEG. We also discussed the mechanisms to calibrate trust and their potential misuse. Finally, we considered the role that visualizations play in increasing trust in AI systems. This includes questions such as: Which mechanisms exist to make AI systems trustworthy? How can interactive visualizations contribute? Under which circumstances are interactive visualizations the decisive factor for enabling responsible AI? And what are the research challenges that still have to be solved – in the area of machine learning or interactive visualization – to leverage this potential in real world applications?

The seminar started with four keynote talks by experts in cognitive psychology, sociology, AI, and visualization, to provide participants with diverse perspectives that helped seed discussion topics. Then, the group decided to build 6 smaller groups to discuss the individual topics that should be worked on during the rest of the week. The six groups collectively came up with a longer list of potential topics surrounding the aspects of trust and machine learning. This list was voted on the plenum to distill it to the following four breakout groups: (1) Good practices and evil knobs in machine learning; (2) Evaluation, measures and metrics for trust in ML; (3) Interaction, expectations and dimension reduction; and (4) Definitions, taxonomy and relationships of trust in ML.

The outcome of this seminar is a better understanding of which aspects of trust have to be considered in fostering trust in AI systems and how interactive visualizations can help foster trust in artificial intelligence systems by making them more understandable and responsible. This will encourage innovative research and help to start joint research projects tackling the issue. Concrete outcomes are drafts of position papers describing the findings of the seminar and in particular, the research challenges identified in the seminar.

6.50 Challenges and Opportunities of Democracy in the Digital Society

Organizers: Abraham Bernstein, Anita Gohdes, Beth Simone Noveck, and Steffen Staab

Seminar No. 22361

Date: September 4–9, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.9.1

Participants: Harith Alani, Pablo Aragón, Fynn Bachmann, Abraham Bernstein, Markus Brill, Anna De Liddo, Martin Emmer, Fabrizio Gilard, Anita Gohdes, Lynda Hardman, Farane Jalali, Eleni Kyza, Jörn Lamla, Libor Pavlícek, Catarina Pereira, Cristina Sarasua, Steffen Staab, Marco Steenbergen, Gefion Thürmer, Miklovana Tuci, Felix Uhlmann

In this Dagstuhl Seminar, we aimed to have interdisciplinary discussions on the *challenges* and *opportunities* of online platforms, online participation, and online deliberation, including experts in politics, law, technology, governance, and policy-making.

In order to achieve a thorough integration of perspectives, we started the first day of the seminar with several keynote talks by scholars from political science, computer science, communication and law. The keynote speakers were Fabrizio Gilardi (Universität Zürich), Anna de Liddo (Open University), Pablo Aragón (Wikimedia Foundation), Eleni Kyza (Cyprus University of Technology), and Felix Uhlmann (Universität Zürich). After these talks, the seminar organized a brainstorming session to identify key discussion topics related to democracy in the digital society. Based on these discussion topics, the participants worked on six breakout sessions: *Goals*, *Actors*, *Narratives and Bias*, *Structure*, *Technology*, and *Success Metrics*. Additionally, throughout the seminar, Markus Brill (Technical University of Berlin), Abraham Bernstein (Universität Zürich), Róbert Bjarnason (Citizens Foundation), Gefion Thürmer (King's College London), Gianluca Demartini (University of Queensland), and Harith Alani (Open University) gave short presentations on various topics including computational social choice, diversity in news recommender systems, citizen science, and misinformation.

The full report provides the abstracts of the talks and the group discussions.

6.51 Model-Driven Engineering of Digital Twins

Organizers: Benoit Combemale, Vinay Kulkarni, Bernhard Rumpe, and Steffen Zschaler

Seminar No. 22362

Date: September 4–9, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.9.20

Participants: Shaukat Ali, Balbir Barn, Ion Barosan, Nelly Bencomo, Francis Bordeleau, Tony Clark, Loek Cleophas, Benoit Combemale, Thomas Godfrey, Georg Grossmann, Gabor Karsai, Djamel Khelladi, Oliver Kopp, Daniel Lehner, Bernhard Mitschang, Paula Muñoz Ariza, Alfonso Pierantonio, Fiona A. C. Polack, Matthias Riebisch, Bernhard Rumpe, Holger Schlingloff, Markus Stumptner, Antonio Vallecillo, Mark van den Brand, Hans Vangheluwe, Andreas Wortmann, Steffen Zschaler

Digital twins are an emerging concept with the potential for revolutionising the way we interact with the physical world. Early versions of digital twins have already been applied successfully in almost all known areas, including engineering and development areas, but also scientific domains and cultural, social, and economic domains.

Digital twins are leveraging the digitalization of increasingly more application domains and are intensively using various models of different types in a descriptive, predictive, and prescriptive way. Furthermore, the the ever-increasing availability of data, the last improvements of sensor technologies, and the reliable connnectivity enable direct inspection and manipulation of real-world systems, both for physical systems and objects as well as social systems and organisations respectively their processes.

The concept has seen strong interest in industry, where there is a desire to control increasingly complex systems of systems, ensuring they behave as expected and to control their adaptation to the environment or any deviations with the initial plan. Digital twins can be used for improved analysis and understanding of complex systems (in silico experimentation) as well as for control and transformation of these systems. Digital twins are themselves complex software systems, posing novel software-engineering challenges, which have so far not been sufficiently addressed by the software-engineering research community.

There is a need for solid foundations to ensure the development of tools and methods according to well-established principles. We believe that Model-Driven Engineering (MDE), will be a key technology for the successful systematic engineering of Digital Twins. In this Dagstuhl Seminar, the goal was to bring together both practitioners and researchers to

- reflect on the concept of Digital Twins and the software-engineering challenges posed,
- identify relevant existing MDE approaches and technologies that can help tackle the challenge of systematically engineering digital twins, and
- define an academia–industry research roadmap for systematic engineering of digital twins based on MDE.

As the intended primary goal of the seminar was to create a community and establish a research roadmap, we have been discussing the following topics:

- Challenges faced in real-world development of Digital Twins.
- Opportunities offered by MDE.
- Active exploration of collaboration opportunities.

The full report reflects the discussions and some of the outcomes, however, we also identified that the overall topic is not only relevant, but also highly innovative, which is why the report does only reflect an intermediate status of discussions and results, but the community will vividly go on to solve the challenges identified and addressed in the rest of the report.

One key outcome of the seminar and its continuing community activities will be to contribute to a solid research roadmap for the new Software Engineering sub-discipline of Model-Based Development of Digital Twins.

Definitions to set the stage

There are two core terms that need appropriate definitions, namely model driven engineering (MDE) and digital twin. While MDE (and with it the terms model and modeling language) are relatively straightforward, there are different variants of definitions for digital twins that served as a starting base.

▶ **Definition** (Model Driven Engineering (MDE)). *Model Driven Engineering* is a state-of-art software engineering approach for supporting the increasingly complex construction and mainte-

nance of large-scale systems [2–4]. In particular, MDE allows domain experts, architects, and developers to build languages and their tools that play an important role in all phases of the development process [7].

As digital twins are currently a relatively young and in particular evolving area of research, it is not surprising, that there is a variety of different definitions available (e.g., [1] identifies more than hundred different definitions).

At the beginning and during the seminar we identified the following definitions to be particularly of interest.

▶ **Definition** (Digital Twin (DT)). A *digital twin* (DT) is a comprehensive digital representation of an actual system, service or product (the Physical Twin, PT), synchronized at a specified frequency and fidelity [5]. The digital twin includes the properties, conditions and behavior of the physical entity through models and data, and is continuously updated with real-time system data [6]. The exchange of data between the digital and the physical twins takes place through bidirectional data connections.

An alternative definition can be found at the website of the software engineering institute of RWTH Aachen University[45] together with some additional discussions:

▶ **Definition** (Digital Twin (DT)). A *digital twin* of a system consists of

- a set of models of the system and
- a set of digital shadows, both of which are purposefully updated on a regular basis,
- provides a set of services to use both purposefully with respect to the original system, and
- can send
  - information about the environment and
  - control commands to the original system.

▶ **Definition** (Digital Twin System (DTS)). Based on these definitions, a *Digital Twin System* can be defined as the combination of an actual system and the DTs of this actual system.

References

1 Manuela Dalibor, Nico Jansen, Bernhard Rumpe, David Schmalzing, Louis Wachtmeister, Manuel Wimmer, and Andreas Wortmann. A cross-domain systematic mapping study on software engineering for Digital Twins. *Journal of Systems and Software*, November 2022.

2 John Hutchinson, Mark Rouncefield, and Jon Whittle. Model-driven engineering practices in industry. In *Proceedings of the 33rd International Conference on Software Engineering*, pages 633–642. ACM, 2011.

3 John Hutchinson, Jon Whittle, Mark Rouncefield, and Steinar Kristoffersen. Empirical assessment of mde in industry. In *Proceedings of the 33rd International Conference on Software Engineering*, pages 471–480. ACM, 2011.

4 Stuart Kent. Model driven engineering. In Michael Butler, Luigia Petre, and Kaisa Sere, editors, *Integrated Formal Methods*, pages 286–298, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.

5 MDT. Digital twin consortium, "glossary of digital twins,". https://www.digitaltwinconsortium.org/glossary/index.html, 2021.

6 Paula Munoz, Javier Troya, and Antonio Vallecillo. Using uml and ocl models to realize high-level digital twins. In *2021 ACM/IEEE International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C)*, pages 212–220. IEEE, 2021.

7 Juha-Pekka Tolvanen and Steven Kelly. Metaedit+: defining and using integrated domain-specific modeling languages. In *The 24th ACM SIGPLAN conference companion on OOPSLA*, pages 819–820, 2009.

[45] https://www.se-rwth.de/essay/Digital-Twin-Definition/

6.52 Algebraic and Analytic Methods in Computational Complexity

Organizers: Markus Bläser, Valentine Kabanets, Ronen Shaltiel, and Jacobo Torán

Seminar No. 22371

Date: September 11–16, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.9.41

Participants: Eric Allender, Markus Bläser, Andrej Bogdanov, Peter Bürgisser, Prerona Chatterjee, Arkadev Chattopadhyay, Gil Cohen, Julian Dörfler, Stephen A. Fenner, Michael A. Forbes, Lance Fortnow, Anna Gál, Frederic Green, Rohit Gurjar, William Hoza, Christian Ikenmeyer, Valentine Kabanets, Pascal Koiran, Antonina Kolokolova, Michal Koucký, Sophie Laplante, Nutan Limaye, Meena Mahajan, Rafael Mendes de Oliveira, Ryan O'Donnell, Natacha Portier, Noga Ron-Zewi, Rahul Santhanam, Nitin Saxena, Rocco Servedio, Ronen Shaltiel, Amir Shpilka, Srikanth Srinivasan, Amnon Ta-Shma, Jacobo Torán, Christopher Umans, Mary Wootters, David Zuckerman, Jeroen Zuiddam

Introduction

The seminar on algebraic methods in computational complexity has traditionally taken place every two years in Dagstuhl for many years. In these meetings, we try to bring together leading researchers in a very active and broad area of theoretical computer science, having the algebraic methods as a unifying thread. Researchers in these areas are relying on ever more sophisticated and specialized mathematics and this workshop can play an important role in educating a diverse community about the latest new techniques, spurring further progress. For the year 2022, we added a new direction that focused besides the algebraic aspect also on methods from analysis. The seminar brought together more than 40 researchers covering a wide spectrum of complexity theory. We had 24 talks, most of them lasting about 45 minutes, leaving ample room for discussions. In the following we describe the major topics of discussion in more detail.

Some areas of focus

Computational complexity is a fundamental and active subarea of theoretical computer science that has produced some of the most well known results in theoretical computer science in recent years. Here we discuss a few broad themes which highlight the importance of algebra as well as analytic methods in computational complexity, and which represent some focus areas of our present seminar.

Circuit complexity

Boolean circuits are one of the most fundamental model of computation. Due to its combinatorial nature, they seem more amenable to formal analysis than the uniform models such as Turing machines. The classical lower bound techniques of Razborov and Smolensky are algebraic: they work by first approximating $\mathrm{AC}^0[p]$ circuits (constant-depth circuits with AND, OR, NOT, and counting modulo prime p gates) by low-degree polynomials, and then proving that certain functions (like Majority) are not well correlated with such polynomials. The Fourier expansion of a Boolean function and its representation as a real multilinear polynomial as well as other analytic tools have been added in the last years to the bag of tools used for the analysis of Boolean circuits. In the seminar, we talked about recent results in circuit complexity.

Andrej Bogdanov talked about property testing. He constructed a natural tester that tells if a function from $\{0,1\}^n$ to some Abelian group is linear (or far from linear).

Frederic Green proved a new correlation bound for certain exponential sums over characteristic 5.

William Hoza presented the construction of a Boolean function F on n bits such that F can be computed by a uniform depth-$(d+1)$ AC^0 circuit with $O(n)$ wires, but F cannot be computed by any depth-d TC^0 circuit with $n^{1+\gamma}$ wires, where $\gamma = 2^{-\Theta(d)}$ and $d = o(\log\log n)$.

Michal Koucký dealt with a classical problem, the simulation of Turing machines by circuits. He gave a new simple proof for the classical result that Turing machines running in time $t(n)$ and space $s(n)$ can be simulated by Boolean circuits of size $O(t(n)\log s(n))$ and of depth $O(t(n))$.

Meena Mahajan presented relations between the minimum rank of a decision tree computing a Boolean function and other complexity measures of the function, as well as a new composition theorem in terms of rank and decision tree depth.

In his talk, *Rocco Servedio* establish a new quantitative version of the Gaussian correlation inequality. It gives a lower bound on the correlation of two centrally symmetric convex sets based on their "common influential directions".

A new family of sampling tasks was presented by *Rahul Santhanam*. He showed that any non-trivial algorithmic solutions

to tasks from this family imply new uniform lower bounds such as "NP not in uniform ACC^0" or "NP does not have uniform depth-2 threshold circuits".

Algebraic complexity

A class of circuits especially suited for the use of algebraic techniques is that of *arithmetic circuits*. These are circuit models that compute polynomial functions by using gates performing arithmetic operations (additions, subtractions, multiplications, divisions, etc.). Two fundamental complexity measures for arithmetic circuits are the *size* and the *depth* or *product depth*.

Prerona Chatterjee considered the question of proving lower bounds against non-commutative circuits better than $\Omega(n \log n)$. She showed a quadratic lower bound against the n-variate central symmetric polynomial.

Arkadev Chattopadhyay talked about connections between communication complexity measures and monotone arithmetic circuit lower bounds. He constructed a (set-multilinear) monotone polynomial that can be computed by depth-3 multilinear formulas in sub-cubic size but requires exponential size to be computed by monotone arithmetic circuits. Second, he proved the existence of a polynomial over n variables in VNP, for which $2^{\Omega(n)}$ size ϵ-sensitive lower bounds hold if $\epsilon = 2^{-O(n)}$.

Barrier results in the group-theoretic approach to bounding the exponent of matrix multiplication was the topic of the talk by *Chris Umans*. He showed that finite groups of Lie type cannot prove $\omega = 2$ and presented a further barrier result. Then he gave constructions in the continuous setting, which can potentially evade these two barriers.

Pascal Koiran studied the decomposition of multivariate polynomials as sums of powers of linear forms. He presented a randomized algorithm for the following problem: Given a homogeneous polynomial of degree d as a blackbox, decide whether it can be written as a linear combination of dth powers of linearly independent complex linear forms.

Nutan Limaye proved in her talk that there exist monomial symmetric polynomials that are hard for the class VNP.

Pseudorandomness and derandomization

The theory of pseudorandomness studies explicit constructions and applications of "random-like" objects of combinatorial or algebraic type. The common feature of such objects is that it is easy to construct one by random sampling, but a very important problem is to get efficient *deterministic* constructions.

Eric Allender proved that Kolmogorov complexity characterizes statistical zero knowledge. Every decidable promise problem has a non-interactive statistical zero-knowledge proof system if and only if it is randomly reducible to a promise problem for Kolmogorov-random strings.

Random walks on expanders are a useful tool in complexity theory. *Gil Cohen* explained how the inherent cost can be reduced from exponential to linear by applying a permutation after each random step.

Sylvester-Gallai type problems have found applications in polynomial identity testing and coding theory. *Rafael Oliveira* discussed such problems and their relation to algebraic computation, and presented a theorem that radical Sylvester-Gallai configurations for cubic polynomials must have small dimension.

Ryan O'Donnell explained how to contruct high-dimensional expanders from Chevalley groups.

Motivated by applications from cryptography, *Noga Ron-Zewi* studied a new interactive variant of PCPs, so-called interactive oracle proofs. She showed that for this model the overhead in the encoding can be made arbitrarily small and the prover complexity overhead can be made constant.

In his talk, *Amon Ta-Shma* gave an alternative construction of the lossless condenser by Guruswami, Umans and Vadhan. Instead of Parvaresh-Vardy codes, the new construction is based on multiplicity codes.

A Chor-Goldreich source is a sequence of random variables where each has min-entropy, even conditioned on the previous ones. *David Zuckerman* showed how to extend this notion in several ways, most notably allowing each random variable to have Shannon entropy conditioned on previous ones. He then proved new pseudorandomness results for Shannon-CG sources.

Border complexity and invariant theory

Many problems in algebraic complexity theory can be written as an orbit closure problem. We are given a vector space V and a group G acting on it. The orbit Gv of an element $v \in V$ is the set $\{gv \mid g \in G\}$ and its closure is the usual closure in the Zariski topology. For instance, we can formulate the tensor border rank problem in this language: Alder and Strassen proved that the question whether a tensor t has border rank $\leq r$ is equivalent to deciding whether t is in the orbit closure (under the standard action $\mathrm{GL}_n \times \mathrm{GL}_n \times \mathrm{GL}_n$) of the so-called unit tensor of size r. As second example is provided by Mulmuley and Sohoni who formulated a variant of the permanent versus determinant question as an orbit closure problem.

Peter Bürgisser gave an introduction to new algorithmic and analysis techniques that extend convex optimization from the classical Euclidean setting to a general geodesic setting. He pointed out the relevance of invariant and representation theory for for complexity theory and highlighted connections to different areas of mathematics, statistics, computer science, and physics.

Rohit Gurjar considered determinants of the matrices of the form $(\sum_i A_i x_i)$ where each A_i is rank one. He showed that this class of polynomials is closed under approximation.

Approximate complexity was also the topic of *Nitin Saxena's* talk. He proved that the border of bounded-top-fanin depth-3 circuits is relatively easy, since it can be computed by a polynomial-size algebraic branching program.

Counting and quantum complexity

In order to study the #P (non-)membership of some concrete problems, *Christian Ikenmeyer* started the development of a classification of the #P closure properties on affine varieties. He obtained oracle separations between counting classes, where the existence of the oracle is based on properties of the vanishing ideal of an affine variety.

Steve Fenner considered a problem in quantum computing, the construction of a "realistic" Hamiltonian for quantum fanout.

Conclusion

The talks of the seminar ranged over a broad assortment of subjects with the underlying theme of using algebraic and analytic techniques. It was a very fruitful meeting and it has hopefully initiated new directions in research. Several participants specifically mentioned that they appreciated the particular focus on a common class of techniques (rather than end results) as a unifying theme of the workshop. We look forward to our next seminar.

6.53 Knowledge Graphs and their Role in the Knowledge Engineering of the 21st Century

Organizers: Paul Groth, Elena Simperl, Marieke van Erp, and Denny Vrandečić

Seminar No. 22372

Date: September 11–14, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.9.60

Participants: Marcel R. Ackermann, Mehwish Alam, Bradley Allen, Sören Auer, Eva Blomqvist, George Fletcher, Paul Groth, Aidan Hogan, Filip Ilievski, Antoine Isaac, Diana Maynard, Deborah L. McGuinness, Axel-Cyrille Ngonga Ngomo, Heiko Paulheim, Lydia Pintscher, Valentina Presutti, Florian Reitz, Marta Sabou, Harald Sack, Stefan Schlobach, Juan F. Sequeda, Elena Simperl, Steffen Staab, Lise Stork, Hideaki Takeda, Katherine Thornton, Marieke van Erp, Denny Vrandečić

Knowledge engineering has changed dramatically in the last twenty years. When the organisers of this seminar were starting out, it used to be about gathering highly curated knowledge from experts and encoding it into computational representations in knowledge bases. It was primarily a manual process, focusing more on how knowledge was structured and organised, for instance, as schemas or ontologies, and less on tying in existing data into that process. The results were used in expert systems and required considerable up-front investment. Today, knowledge base construction is a largely automatic process with human-in-the-loop. Owing to greater availability of data in different modalities and to advances in data management, machine learning, and crowdsourcing, knowledge bases today incorporate large amounts of knowledge. Provided access to data and (off-the-shelf) AI capabilities, an organisation can create a large knowledge base at a fraction of the costs from decades ago. It's for these reasons that we see knowledge bases, in particular in the form of knowledge graphs, routinely applied in anything from search and intelligent assistants to digital twins, supply chain management, and legal compliance. Many socio-technical challenges remain, which the seminar aimed to address with a mix of invited talks, deep-dives, and small-group workshops as following:

Landscape review: as the field has changed so much, both in research and practices, it was important to take inventory of approaches, methods, techniques, and tools by analysing real-world case studies where knowledge bases and knowledge graphs are created and used. Participants reflected on core lessons learned, knowledge gaps, and opportunities to create and maintain knowledge graphs at scale in various domains.

The knowledge graph life cycle: participants discussed extant knowledge engineering pipelines and identified gaps and connections between knowledge sources and methods and tools used in the construction and maintenance of knowledge graphs, including large language models and generative AI systems. There was consensus that we need a sustained effort to update and upgrade classical ontology engineering methodologies and develop a prototype infrastructure to make the most of the latest neurosymbolic technologies and tools. One specific challenge identified during the seminar was around taking knowledge engineering and knowledge graphs beyond structured data e.g., tables and information extraction from text to other modalities.

Using AI responsibly: as knowledge graph construction is slowly but surely embracing more and more sophisticated AI capabilities to scale, it is critical that processes and outcomes are aligned with fairness, accountability, and transparency guidance and standards. Solutions need to consider a range of end-users and stakeholders, including those that are unique to knowledge engineering settings such as domain experts, information scientists and librarians, and knowledge graph developers. Participants discussed the need for setting up task-based studies and in-depth analyses of human-centric challenges, and for developing bespoke explainability solutions and bias and fairness assessments.

Knowledge and technology transfer: knowledge graphs and knowledge engineering do not exist in isolation. From a research point of view, participants suggested activities to build capabilities to use the latest neurosymbolic technologies and tools in knowledge graph construction, including tutorials, workshops, and hackathons, and to jointly develop frameworks and methodologies. From an application point of view, it was recognised that there is a need to promote knowledge graphs to the wider developer community and communicate their benefits, for instance, alongside neural methods.

6.54 Rational Design of RiboNucleic Acids

Organizers: Sven Findeiß, Christoph Flamm, and Yann Ponty

Seminar No. 22381

Date: September 18–23, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.9.121

Participants: Ebbe Sloth Andersen, Maciej Antczak, Stefan Badelt, Danny Barash, Sarah Berkemer, Anne Condon, Harold Fellermann, Jorge Fernández de Cossío Díaz, Sven Findeiß, Christoph Flamm, Cody Geary, Jan Gorodkin, Christine Heitsch, Ivo Hofacker, Felix Kühnl, István Miklós, Philippe Nghe, Cyrille Merleau Nono Saha, Samuela Pasquali, Katja Petzold, Yann Ponty, Vladimir Reinharz, Lorenz Ronny, Frederic Runge, Bruno Sargueil, Nicolas Schabanel, Shinnosuke Seki, Petr Sulc, Marta Szachniuk, Andrew Torda, Maria Waldl, Sebastian Will, Hua-Ting Yao

Context and selected takeaways

RiboNucleic Acids (RNAs) are ubiquitous macromolecules within biological systems, capable of performing a wide range of regulatory and catalytic functions. This versatility can be harnessed, and RNAs are increasingly utilized to accurately monitor and control biological processes [19], leading to RNA being found at the core of modern therapeutics [18]. It is therefore not surprising that the RNA-guided CRISPR-Cas9 editing [10], rewarded by the 2020 Nobel Prize in Chemistry, and mRNA-based vaccines [12], are at the forefront of modern biotechnology. For many functional RNA families [11], decades of research have produced a deep understanding of the sequence and structural basis underlying their biological function(s). Such studies, coupled with mature computational methods for structure prediction [23], have paved the way for a rational design of RNAs targeting a wide diversity of biological function [2, 8, 13].

Accordingly, RNA design has emerged as an exciting open computational problems in molecular biology. Owing to the discrete nature of RNA sequence and popular structural representations (e.g., secondary structure), RNA design has inspired the contribution of a large number of diverse algorithms [4, 9, 14, 20] for the inverse folding problem, i.e., the design of an RNA sequence which preferentially and effectively folds into a predefined (secondary) structure. Given the, recently established, NP-Hardness of the problem, even for minimal energy models [1], many of those algorithmic predictions are either heuristics, exponential-time or based on a variety of machine learning techniques.

More generally, RNA design addresses the generation of sequences of nucleotides targeting a given biological function. A non-exhaustive list of classic design objectives includes:

- Preferential adoption of one or multiple given structures (inverse folding);
- Sequence specific constraints such as an overall (di-)nucleotide composition [21], encoding of an amino-acid sequence (mRNA design), presence/absence of motifs [22];
- Adoption of different conformations upon presence of ligand (RNA switches and sensors) [3];
- Effective and specific interactions with targeted partners (RNAs, proteins) cascading into system-level regulatory effects [15, 16];
- Self-assembly into large scale architectures, ultimately adopting a predefined 3D shape (RNA origami) [6];
- Exploit co-transcriptional folding, and more general out-of-equilibrium regimes to perform computations (strand displacement systems, oritatami) [5]

Typical applications of design include novel therapeutic strategies, control principles for existing biological systems, or sensors for the presence of small molecules [3], but designed sequences can also provide an objective experimental assessment of functional hypotheses, where designs are synthesized and their effect on the cellular context can be tested *in vitro* and, in turn, *in vivo*.

Over the course of the seminar, we witnessed a substantial recent expansion of the scope of applications. Beyond classic but still challenging objectives of design, including riboswitches addressed by Talk 5.8[46], 5.9 and 5.21, messenger RNAs towards

[46] Here and in the following, talk numbers refer to the section in the full report documenting the talk.

vaccine objectives mentioned by Talk 5.27, and CRISPR gRNAs mentioned by Talk 5.11, novel applications of RNA design emerged during the seminar. Talk 5.11 introduced SARS-CoV-2 sensors based on strand displacement, Talk 5.15 addressed self-replicating ribozymes connected with origin-or-life questions, and Talk 5.6 explored rational design principles for repetitive RNAs inducing the formation of cellular droplets through liquid-liquid phase separation.

RNA Design as a discrete (inverse) optimization problem The inverse folding problem, one of the central elements of RNA design, is a hard computational problem [1]. Although attracting a wide interest from the community, it is also one of the very few problems in computational biology whose complexity status has remained open for a long time (about three decades). This difficulty can be attributed to a lack of a suitable conceptual framework for inverse combinatorial problems. Indeed, inverse folding can be viewed as the search of a pre-image, in a function that maps each RNA sequence to its most stable conformation, the latter being computed using a polynomial, yet non-trivial, dynamic programming algorithm [23]. Natural generalizations virtually include any instance of inverse optimization problems, and could be of general interest to the Computer Science research field. Prior works in this direction have led to characterization of designable structures based on formal languages and graph theory [7], revealing strong connections to many subfields of computer science (for instance, between positive design and graph coloring).

In Talk 5.18 it was discussed that a flexible inverse folding approach, e.g., by allowing the extension of helices by at most one base pair, seems to be easier than keeping the problem strict. Such a flexibility in the structural objective of design was also emphasized as desirable by Talk 5.5. The problem of classical inverse folding can be extended from one to multiple target structures, and Talk 5.27 showed that this can be solved by an elegant dynamic programming approach that is fixed parameter tractable. The resulting framework was further generalized, and is not only applicable to RNA design, but also to apparently more distant problems such as the alignment of RNAs with pseudoknots. *In silico* designs and analysis depend on the accuracy of the applied energy model. In Talk 5.12 it has been underpinned that a systematic perturbation of parameters can be used to define a notion of robustness of individual parameters of an energy model, and help to improve prediction accuracy. Talk 5.28 revisited the inverse folding problem as an inverse optimization problem, and showed that many local structural motifs do not admit a design, with consequences to the space of designable structures, but raised fundamental questions on a relatively new flavor of optimization.

RNA Design in Structural Bioinformatics Inverse folding also represents the ultimate test of our understanding of the mechanisms governing the folding of macromolecules. Given a set of folding rules (typically, an energy model), a synthesis of *in silico* designed sequences combined with high-throughput experiments (e.g., structure probing) enables an assessment of the compatibility of the determined structure with the initial target. Observed discrepancies can then be used to assess the quality of predictive models, especially those based on statistical potentials which may be prone to overfitting. Systematic local imprecisions can also be used to refine energy models, enabling the generation of better designs, whose iteration represents a virtuous circle, ultimately contributing to a better understanding of folding principles.

A nascent RNA molecule typically folds during its transcription. Frameworks to simulate this kinetically driven process can help to interpret experimental results (Talk 5.4) but as neither the simulation nor the experiment is perfect, quality assurance (Talk 5.14) of the *in silico* investigations is essential and results have to be interpret with caution, as for instance the mapping of time scales is a non trivial task. Finally, complex RNA hybridization networks are designed *in silico* to perform regulatory functions with complex temporal dynamics. A simplified kinetic model, introduced in Talk 5.26, for RNA/RNA hybridization represents an attractive evaluation model for the design of interactions.

At a much more detailed 3D level, Talk 5.17 showed that high-resolution experimental techniques can be used to observe dynamic behaviors, sometimes triggered by the binding of a ligand, and could inform future objective functions. Talk 5.24 and Talk 5.16 described coarse-grained models amenable to molecular dynamics. Interestingly, the latter can be leveraged in order to study kinetics behaviors at the 3D level.

RNA Design in Synthetic Biology and Natural Computing This line of research applies various engineering principles to the design, and construction of artificial biological devices. While initially focused on hijacking naturally-occurring regulatory functions through a copy/paste of evolved genetic material [17], the need for a precise control and for a modularity/orthogonality of constructs, has increasingly led to a *de novo* design based on nucleic acids. Recently, RNA has been successfully used as a material for the design of whole regulatory circuits, or for the construction of complex programmable shapes (RNA origamis [6]), with promising applications as biomaterial.

Software frameworks, like the ones presented in Talk 5.22 and Talk 5.10, make the construction of large DNA and RNA nano-structures possible. Those designs are not only adopting the right structure *in silico* according to combinatorial folding algorithms, but can also be validated by simulations (Talk 5.24) and microscopy (Talk 5.1). This observation suggests that the difficulty of design could stem from the compactness of targeted ncRNAs, while larger (but more regular) RNAs may be easier to design, an element that could inspire future theoretical studies.

In the course of the seminar it became evident that information from the 2D and 3D level need to be mapped onto each other. Design could therefore benefit from multiscale approaches: selecting candidates with 2D objectives, use coarse grained 3D analysis (Talk 5.16) and go to a full atom final validation for critical sub-regions. The curation of refined and non-redundant 3D RNA structures (Talk 5.2) and the systematic extraction of information from such a data set can help to investigate for instance structural features of modified bases or to propose isosteric structural mutations (Talk 5.19) in order to generalize the design from 2D to compact 3D architectures.

Programmable RNA folding can also be used as a computational model, allowing for the computation of complex programs based on cotranscriptional folding phenomena. RNA regulatory circuits can be used to emulate Boolean functions, allowing a precise and expressive control of regulatory networks at an early stage of the gene expression process. Talk 5.23 introduced RNA oritatami, a Turing-complete model of computation based on cotranscriptional folding inspired by cellular automata. Talk 5.7 described exciting applications of design to generate easily-checkable QR codes that reveal contamination in a closed environment. However an application-agnostic implementation of the strand displacement systems underlying some of those applications still represent major challenges in RNA, as shown and discussed during Talk 5.3. Those include intra-molecular base pairs and

an overall wasteful behavior that motivates efficient recycling strategies.

RNA evolution and Machine Learning The analysis of new RNA families, such as the pervasive and poorly understood lncRNAs or the numerous viral/bacterial non-coding RNAs observed in metagenomics experiments, relies critically on the identification of an evolutionary pressure, allowing to hypothesize new functions. Given a family of homologous RNAs sharing established functional traits, it is classic to asks whether an observed property, such as the occurrence of a common motif or a given covariation pattern, is likely to reveal a yet-unknown selective pressure or, conversely, is merely the consequence of established functional traits. Classic bioinformatics methods rephrase the problem in a hypothesis-testing framework, and compute the probability that a sequence, generated at random in a model that captures existing constraints, features the observed property. Ideally, such sequences should represent solutions to an instance of the design problem, target established functions, while respecting a distribution that can either be derived from the targeted function, or learned from data.

Talk 5.5 presented a context where rational design methodologies were utilized to capture remote homologs of a suspected, but scarcely-populated, functional family of ncRNAs. Generative models can also be used for design, in cases where the underlying model of function is partially understood, and should be learned from the data. Talk 5.8 used Restricted Boltzmann Machines (RBM), an unsupervised learning approach, to pickup the intricate probability distribution of the statistical features of a naturally-occurring riboswitch. The RBM was then used to generate novel instances with the same distribution of features, resulting in an enrichment of functional designs as revealed by experimental validation. Direct Coupling Analysis was also used in Talk 5.15 to generate self-replicating ribozymes, using a complex definition of function that may require some element of learning. Interestingly, the efficacy of designs was ultimately shown to benefit from further refinements using classic combinatorial methods for inverse folding, suggesting future hybrid ML/combinatorial methodologies.

While RNA design is an increasingly important computational task in molecular biology, nanotechnology and medicine, methods for computational rational design are still lacking for many applications. Moreover, many design tasks are currently addressed using algorithmic techniques (e.g., Markov chain Monte Carlo) that are clearly superseded by the state-of-the-art in algorithmic research. Conversely, computer scientists considering design tasks usually limit themselves to inverse folding, overlooking a rich bestiary of computational problems whose consideration would, in turn, undoubtedly lead to the emergence of new algorithmic paradigms.

Talk 5.20 showed that basic ML architectures can be learned in the context of reinforcement learning and can be successful for basic inverse folding of RNA. Talk 5.25 presented a complete design story, describing a methodology to advance our understanding of tRNAs. In particular, a mechanistic understanding of the target function can be gained by masking constraints during redesign, and the differentiability of the design problem can lead to great speedups of the computation. However, ML approaches may not always represent a silver bullet in RNA bioinformatics, and Talk 5.13 dramatically illustrated this in the context of RNA folding, a context where the quality and biases in the data strongly impacts, and probably hinders for intrinsic reasons, the predictive capabilities of deep learning-based methods. As a consequence, synthetic data can and should be used to test the capacity of learning architectures on simplified problems before embarking into “real life” learning.

References

1 Édouard Bonnet, Paweł Rzążewski, and Florian Sikora. Designing RNA secondary structures is hard. In Benjamin J. Raphael, editor, *Research in Computational Molecular Biology – 22nd Annual International Conference, RECOMB 2018*, volume 10812 of *Lecture Notes in Computer Science*, pages 248–250, Paris, 2018. Springer.

2 Matthew G. Costales, Jessica L. Childs-Disney, Hafeez S. Haniff, and Matthew D. Disney. How we think about targeting RNA with small molecules. *Journal of Medicinal Chemistry*, 63(17):8880–8900, 2020. PMID: 32212706.

3 Sven Findeiß, Maja Etzel, Sebastian Will, Mario Mörl, and Peter F Stadler. Design of artificial riboswitches as biosensors. *Sensors (Basel, Switzerland)*, 17(9):E1990, August 2017.

4 Juan Antonio Garcia-Martin, Peter Clote, and Ivan Dotu. RNAiFOLD: a constraint programming algorithm for RNA inverse folding and molecular design. *Journal of Bioinformatics and Computational Biology*, 11(2):1350001, 2013.

5 Cody Geary, Pierre-Étienne Meunier, Nicolas Schabanel, and Shinnosuke Seki. Oritatami: A Computational Model for Molecular Co-Transcriptional Folding. *International Journal of Molecular Sciences*, 20(9):2259, May 2019.

6 Cody Geary, Paul W. K. Rothemund, and Ebbe S. Andersen. A single-stranded architecture for cotranscriptional folding of RNA nanostructures. *Science*, 345(6198):799–804, 2014.

7 Jozef Haleš, Alice Héliou, Ján Maňuch, Yann Ponty, and Ladislav Stacho. Combinatorial RNA Design: Designability and Structure-Approximating Algorithm in Watson-Crick and Nussinov-Jacobson Energy Models. *Algorithmica*, 79(3):835–856, Nov 2017.

8 Stefan Hammer, Christian Günzel, Mario Mörl, and Sven Findeiß. Evolving methods for rational de novo design of functional RNA molecules. *Methods*, 161:54–63, may 2019.

9 I.L. Hofacker, W. Fontana, P.F. Stadler, L.S. Bonhoeffer, M. Tacker, and P. Schuster. Fast folding and comparison of RNA secondary structures. *Monatsch. Chem.*, 125:167–188, 1994.

10 Martin Jinek, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, 337(6096):816–821, 2012.

11 Ioanna Kalvari, Joanna Argasinska, Natalia Quinones-Olvera, Eric P Nawrocki, Elena Rivas, Sean R Eddy, Alex Bateman, Robert D Finn, and Anton I Petrov. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic acids research*, 46:D335–D342, January 2018.

12 David M. Mauger, B. Joseph Cabral, Vladimir Presnyak, Stephen V. Su, David W. Reid, Brooke Goodman, Kristian Link, Nikhil Khatwani, John Reynders, Melissa J. Moore, and Iain J. McFadyen. mRNA structure regulates protein expression through changes in functional half-life. *Proceedings of the National Academy of Sciences*, 116(48):24075–24083, 2019.

13 Na Qu, Yachen Ying, Jinshan Qin, and Antony K Chen. Rational design of self-assembled RNA nanostructures for HIV-1 virus assembly blockade. *Nucleic Acids Research*, 50(8):e44–e44, 12 2021.

14 Vladimir Reinharz, Yann Ponty, and Jérôme Waldispühl. A weighted sampling algorithm for the design of RNA sequences with targeted secondary structure and nucleotide distribution. *Bioinformatics*, 29(13):i308–i315, 2013.

15 Guillermo Rodrigo, Thomas E Landrain, and Alfonso Jaramillo. De novo automated design of small RNA circuits for engineering synthetic riboregulation in living cells. *Proceedings of the National Academy of Sciences U S A*, 109(38):15271–6, 2012.

16 Guillermo Rodrigo, Thomas E. Landrain, Eszter Majer, José-Antonio Daròs, and Alfonso Jaramillo. Full design automation of multi-state RNA devices to program gene expression using energy-based optimization. *PLoS Computational Biology*, 9(8):e1003172, 2013.

17 E Westhof, B Masquida, and L Jaeger. RNA tectonics: towards RNA design. *Folding & design*, 1(4):R78—88, 1996.

18 Melanie Winkle, Sherien M. El-Daly, Muller Fabbri, and George A. Calin. Noncoding RNA therapeutics – challenges and potential solutions. *Nat Rev Drug Discov*, 20:629–651, 2021.

19 Sherry Y. Wu, Gabriel Lopez-Berestein, George A. Calin, and Anil K. Sood. RNAi therapies: Drugging the undruggable. *Science Translational Medicine*, 6(240):240ps7, 2014.

20 Joseph N Zadeh, Brian R Wolfe, and Niles A Pierce. Nucleic acid sequence design via efficient ensemble defect optimization. *Journal of Computational Chemistry*, 32(3):439–52, 2011.

21 Yang Zhang, Yann Ponty, Mathieu Blanchette, Eric Lecuyer, and Jérôme Waldispühl. SPARCS: a web server to analyze (un)structured regions in coding RNA sequences. *Nucleic Acids Research*, 41(Web Server issue):W480–5, July 2013.

22 Yu Zhou, Yann Ponty, Stéphane Vialette, Jérôme Waldispühl, Yi Zhang, and Alain Denise. Flexible RNA design under structure and sequence constraints using formal languages. In Jing Gao, editor, *ACM Conference on Bioinformatics, Computational Biology and Biomedical Informatics. ACM-BCB 2013, Washington, DC, USA, September 22–25, 2013*, page 229. ACM, 2013.

23 M. Zuker and P. Stiegler. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Research*, 9:133–148, 1981.

6.55 Machine Learning for Science: Bridging Data-Driven and Mechanistic Modelling

Organizers: Philipp Berens, Kyle Cranmer, Neil D. Lawrence, Jessica Montgomery, and Ulrike von Luxburg

Seminar No. 22382

Date: September 18–23, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.9.150

Participants: Mauricio A Álvarez, Bubacarr Bah, Jessica Beasley, Philipp Berens, Maren Büttner, Kyle Cranmer, Thomas G. Dietterich, Carl Henrik Ek, Stuart Feldman, Asja Fischer, Philipp Hennig, David W. Hogg, Christian Igel, Samuel Kaski, Ieva Kazlauskaite, Hans Kersting, Niki Kilbertus, Vidhi Lalchand, Neil D. Lawrence, Gilles Louppe, Dina Machuve, Jakob Macke, Eric Meissner, Siddharth Mishra-Sharma, Jessica Montgomery, Jonas Peters, Aditya Ravuri, Markus Reichstein, Bernhard Schölkopf, Francisco Vargas, Soledad Villar, Ulrike von Luxburg, Verena Wolf

Today's scientific challenges are characterised by complexity. Interconnected natural, technological, and human systems are influenced by forces acting across time- and spatial-scales, resulting in complex interactions and emergent behaviours. Understanding these phenomena – and leveraging scientific advances to deliver innovative solutions to improve society's health, wealth, and well-being – requires new ways of analysing complex systems.

Artificial intelligence (AI) offers a set of tools to help make sense of this complexity. In an environment where more data is available from more sources than ever before – and at scales from the atomic to the astronomical – the analytical tools provided by recent advances in AI could play an important role in unlocking a new wave of research and innovation. The term AI today describes a collection of tools and methods, which replicate aspects of intelligence in computer systems. Many recent advances in the field stem from progress in machine learning, an approach to AI in which computer systems learn how to perform a task, based on data.

Signals of the potential for AI in science can already be seen in many domains. AI has been deployed in climate science to investigate how Earth's systems are responding to climate change; in agricultural science to monitor animal health; in development studies, to support communities to manage local resources more effectively; in astrophysics to understand the properties of black holes, dark matter, and exoplanets; and in developmental biology to map pathways of cellular development from genes to organs. These successes illustrate the wider advances that AI could enable in science. In so doing, these applications also offer insights into the science of AI, suggesting pathways to understand the nature of intelligence and the learning strategies that can deliver intelligent behaviour in computer systems.

Further progress will require a new generation of AI models. AI for science calls for modelling approaches that can: facilitate sophisticated simulations of natural, physical, or social systems, enabling researchers to use data to interrogate the forces that shape such systems; untangle complicated cause-effect relationships by combining the ability to learn from data with structured knowledge of the world; and work adaptively with domain experts, assisting them in the lab and connecting data-derived insights to pre-existing domain knowledge. Creating these models will disrupt traditional divides between disciplines and between data-driven and mechanistic modelling.

The roadmap presented here suggests how these different communities can collaborate to deliver a new wave of progress in AI and its application for scientific discovery. By coalescing around the shared challenges for AI in science, the research community can accelerate technical progress, while deploying tools that tackle real-world challenges. By creating user-friendly toolkits, and implementing best practices in software and data engineering, researchers can support wider adoption of effective AI methods. By investing in people working at the interface of AI and science – through skills-building, convening, and support for interdisciplinary collaborations – research institutions can encourage talented researchers to develop and adopt new AI for science methods. By contributing to a community of research and practice, individual researchers and institutions can help share insights and expand the pool of researchers working at the interface of AI and science. Together, these actions can drive a paradigm shift in science, enabling progress in AI and unlocking a new wave of AI-enabled innovations.

The transformative potential of AI stems from its widespread applicability across disciplines, and will only be achieved through integration across research domains. AI for science is a rendezvous point. It brings together expertise from AI and application domains; combines modelling knowledge with engineering know-how; and relies on collaboration across disciplines and between humans and machines. Alongside technical advances, the next wave of progress in the field will come from building a community of machine learning researchers, domain experts, citizen scientists, and engineers working together to design and deploy effective AI tools.

6.56 Cognitive Robotics

Organizers: Fredrik Heintz, Gerhard Lakemeyer, and Sheila McIlraith

Seminar No. 22391

Date: September 25–30, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.9.200

Participants: Michael Beetz, Mohamed Behery, Jens Claßen, Anthony Cohn, Frank Dignum, Alexander Ferrein, Michael Fisher, Hector Geffner, Jasmin Grosinger, Nick Hawes, Fredrik Heintz, Till Hofmann, Mikhail Khodak, Sven Koenig, Gerhard Lakemeyer, Yves Lesperance, Setareh Maghsudi, Cynthia Matuszek, Sheila McIlraith, Christian Muise, Bernhard Nebel, Tim Niemueller, Ron Petrick, Sebastian Sardiña, Matthias Scheutz, Stefan Schiffer, Maayan Shvo, Gerald Steinbauer, Brian C. Williams

Cognitive Robotics is concerned with endowing robots or software agents with higher level cognitive functions that involve reasoning, for example, about goals, perception, actions, the mental states of other agents, collaborative task execution, etc. This research agenda has historically been pursued by describing, in a language suitable for automated reasoning, enough of the properties of the robot, its abilities, and its environment, to permit it to make high-level decisions about how to act. Such properties were typically encoded by a human, but with recent advances in machine learning, many of these properties, and the determination of how to act, can be learned or adapted through experience. This in turn raises the question of how we can ensure that robots, or other intelligent agents, can be constructed in a manner that is compatible with human values and modes of interactions.

The Cognitive Robotics workshop series has been running since 1998 and includes a Dagstuhl Seminar held in 2010. While progress in Cognitive Robotics has undoubtedly been made over the past twenty years, it is fair to say that we are still far away from creating truly cognitive robots. In particular, the years since the previous Dagstuhl Seminar have seen tremendous progress in many areas that touch on the realisation of cognitive robots such as advances in human-robot interaction and machine learning.

This seminar featured sessions devoted to the following four themes:

Cognitive Robotics and KR: While knowledge representation and reasoning (KR) has played a role in robotic systems for many years, for example, by incorporating domain knowledge in the form of description logic-based ontologies or using automated planning systems for high-level robot control, obstacles remain, which prevent today's robots from benefiting from the true potential of KR. In this session we re-visited the state of the art of how KR is used in robotics and discussed challenges and possible benchmark problems that would demonstrate the need and benefit of KR techniques for cognitive robots. The session was organized by Michael Beetz, University of Bremen.

Verification of Cognitive Robots: Verification has been an active research area in formal methods for many years. It is also an important topic when it comes to cognitive robots, especially when it comes to achieving trustworthiness. However, the sheer complexity of the interplay between a robot's hard- and software components makes verification particularly challenging. In this session we discussed where we currently stand in terms of verifying cognitive robots and what challenges lie ahead. The session was organized by Michael Fisher, University of Manchester.

Human-Robot Interaction and Robot Ethics: For cognitive robots to be useful in human environments, effective human-robot interaction (HRI) plays a crucial role. Besides the technological challenges such as multi-modal communication, ethical considerations have become more and more important. These range from robots observing norms and conventions to humans viewing robots as moral agents. In this session we discussed the many facets of robot ethics in the context of HRI and identified a number of future challenges and open problems. The session was organized by Matthias Scheutz, Tufts University.

Planning and Learning: While planning and learning have traditionally been separate research tracks in cognitive robotics, recent work has shown how action primitives that form the basis of planning can be learned from data without background knowledge, thus avoiding the need for hand-crafted solutions. In this session this work and related proposals were discussed and a roadmap with short- and long-term challenges was drawn up. The session was organized by Hector Geffner, ICREA and Universitat Pompeu Fabra, Spain.

The format of the sessions varied and consisted of one or more plenary talks, plenary discussions, and/or working groups. Working groups for all four themes discussed challenges and roadmaps for the future, and one representative of each group presented their findings on the last day of the seminar. Besides talks and discussions that centered around the four themes, the seminar also featured two invited talks by Luis Lamb, Universidade Federal Do Rio Grande Do Sul, on neurosymbolic AI and by Jan Peters, TU Darmstadt, on robot learning. In addition, a number of participants gave poster presentations on their research.

The organizers of the seminar wish to thank Schloss Dagstuhl for providing such an excellent environment for exchanging ideas on how to move the field of cognitive robotics forward.

6.57 Transparent Quantitative Research as a User Interface Problem

Organizers: Kasper Hornbæk, Yvonne Jansen, Amelia A. McNamara, Judy Robertson, and Chat Wacharamanotham

Seminar No. 22392

Date: September 25–30, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.9.220

Participants: Lonni Besancon, Sophia Crüwell, Pierre Dragicevic, Julien Gori, Lahari Goswami, Lynda Hardman, Olga Iarygina, Yvonne Jansen, Eunice Jun, Ulrik Lyngs, Amelia A. McNamara, Duong Nhu, Viktorija Paneva, Michael Sedlmair, Kavous Selahzadeh Niksirat, Theophanis Tsandilas, Jan Benjamin Vornhagen, Chat Wacharamanotham, Erich Weichselgartner, Wesley J. Willett

Introduction

Many scientific fields face a **replication crisis**: A sizable portion of quantitative research studies could not be replicated. When these studies were re-run with higher statistical power (i.e., more participants), their results yielded effects substantially weaker or even opposite of that in the original studies. This lack of replicability threatens the credibility of research claims and undermines the general public's trust in science. The replication crisis motivated the **Open Science** movement that **promotes transparency throughout the scientific process**: research funding, research design, data collection and analysis, peer reviewing, and knowledge dissemination. These phenomena attracted the interest of researchers in the fields of **Human–Computer Interaction (HCI)** and **Visualization (VIS)** for two reasons. Like other fields, HCI and VIS researchers face challenges in promoting transparency among their peers, effectively implementing and educating transparent practices, and incorporating transparency in the research evaluation processes. However, HCI and VIS researchers have the methods and skills to empirically study these phenomena and design potential solutions. The fields of HCI and VIS also provide a challenging testbed for these inventions.

This Dagstuhl Seminar initiated and advanced works on these issues by bringing together 23 researchers from HCI, VIS, statistics, psychology, data science, and philosophy. They were from Australia, Austria, Canada, Denmark, Finland, France, Germany, the Netherlands, Sweden, Switzerland, the UK, and the USA. Three participants joined online due to the COVID situation and travel difficulties.

Program

We worked in groups to identify problem areas and prototype potential solutions in a Hackathon. We solicited feedback on these prototypes from conference and journal editors and community leaders. The seminar unfolded as follows:

Day 1: After a brief introduction to the purpose of the seminar and the overall plan, participants discussed in small groups to identify problems and challenges to work on in the Hackathon. These discussions were intentionally designed to be free-form to avoid prematurely limiting the areas of interest. To stimulate discussions and spark ideas, we provided the participants access to free-text responses to a survey on the perception of research transparency that we collected from HCI researchers in the weeks before the seminar. Four rounds of discussion were interleaved with three-minute presentations of intermediate results in the plenary to facilitate convergence and consolidation.

In each plenary round, we also asked a few participants to interview each other in front of the room to acquaint everyone with their background and research interest. Day 1 concluded with four clusters of topics to be worked on: (1) Educating researchers, (2) Clarifying the threats from the lack of transparency, (3) Clarifying the "transparency" concept, and (4) Working on how to influence policy and procedures in the publication process.

Day 2: Participants joined the problem cluster according to their interests and started the Hackathon. We provided each group with collaborative workspaces on Google Docs and Miro (an online whiteboard platform). After two Hackathon sessions in the morning, we further stimulated their work with an input lecture from Tim Errington, the Senior Director of Research

at the Center for Open Science (see the full version for the abstract). This lecture highlighted challenges in promoting research transparency and provided a framework for changing research culture at multiple levels: from top-down research funding policy and bottom-up to ease the implementation of transparent practices by providing infrastructure and incentives. After the lecture, the Hackathon continued. We wrapped up the day with a 3-minute presentation from each group and a plenary discussion.

Day 3: The Hackathon continued in the morning. We gave the participants prompts to encourage them to hone in on a concrete idea and realize a prototype that demonstrates the idea's essence. The afternoon is free time for the participants to self-organize group activities to promote trust and informal interactions. We did not organize an excursion because the transportation companies were unavailable.

Day 4: The Hackathon continued in the morning. In the afternoon, the participants presented their preliminary results to four panelists who joined online. The panelists hold influential positions in the research publication process in HCI and VIS: the SIGCHI President, the TOCHI editor-in-chief, the TVCG Associate Editor in Chief and Eurographics Publication Board, CG&A Associate Editor-in-chief, TVCG Associate Editor, and the vice-chair of the IEEE VIS Steering Committee. A discussion on feedback from policy-making perspectives followed each presentation. The conversation with the panelists broadened participants' views about stakeholders and potential concerns. After the discussion, there was a plenary discussion to process the input from the panel collectively. We identified four areas to work on in the manifesto: definition, benefits, subfield-suitability, and progressive transparency.

Day 5: Participants worked in groups to draft a manifesto on research transparency. The seminar concluded with a plenary session where we identified possible future projects, their follow-up actions, and coordinators.

Results

The tangible results of the seminar comprise four prototypes from the Hackathon and a draft manifesto:

1. To influence research funders' policies, we drafted a list of policy suggestions for incentivizing research transparency and Open Science.
2. To inspire researchers and students, we prototyped how we could collect, analyze, and showcase papers in the visualization field that are exemplary in their transparent practices.
3. To improve infrastructure, we identified low-hanging fruits in improving the user interfaces of the ethical review and publication processes to encourage transparent practices.
4. To ease the adoption of transparency practices, we prototyped a cheat sheet that provides reminders for considering transparent practices at each of the research stages. The cheat sheet also provides pointers to relevant guides and resources.

The draft manifesto clarifies the definition of research transparency, describes its benefit, calls for each subfield to identify its suitable set of transparent practices, and argues how transparent practices could be viewed as a progression instead of demanding everyone to be perfect at their first try. These results provide a starting point for future follow-up research and educational activities that will advance the understanding and adoption of research transparency in HCI, VIS, and beyond.

6.58 Computer Science Methods for Effective and Sustainable Simulation Studies

Organizers: Wentong Cai, Christopher Carothers, David M. Nicol, and Adelinde M. Uhrmacher

Seminar No. 22401

Date: October 3–7, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.10.1

Participants: Philipp Andelfinger, Luca Bortolussi, Wentong Cai, Christopher Carothers, Rodrigo Castro, Joachim Denil, Jérôme Feret, Peter Frazier, Reiner Hähnle, Dong (Kevin) Jin, Franziska Klügl, Till Köster, Michael Lees, Jason Liu, Margaret Loper, Fabian Lorig, Bertram Ludäscher, Kresimir Matkovic, Laura Nenzi, David M. Nicol, Alessandro Pellegrini, Niki Popper, Caitlin Ross, Cristina Ruiz-Martin, Bernhard Rumpe, Susan Sanchez, Nadja Schlungbaum, Peter Sloot, Claudia Szabo, Wen Jun Tan, Adelinde M. Uhrmacher, Gabriel A. Wainer, Pia Wilsdorf, Verena Wolf

Motivation. Simulation becomes more and more important in application areas, establishing itself as the third way of science in addition to theory and (real) experiments. To answer research questions, simulation studies form increasingly intricate processes that intertwine the design and execution of various, often calculation-intensive simulation experiments, the generation and refinement of simulation models, and steps of analysis.

This Dagstuhl Seminar has been dedicated to addressing central methodological challenges in supporting the conduction of effective and sustainable simulation studies. Thereby, the seminar focused on problems and solutions related to improving:

- Effectiveness: the usage of resources, including computing infrastructure and data, and the assistance of humans throughout a simulation study.
- Sustainability: continuing a simulation study into the future through support for reusing or building upon its central products, such as simulation model, data, and processes as well as the software used.

The last decades have seen a wide range of methodological developments in computer science that are likely to be instrumental in achieving effective and sustainable simulation studies. However, those efforts are scattered across different computer science fields that include high-performance computing, (modeling) language design, operations research, visual analytics, workflows, provenance, and machine learning, as well as modeling and simulation. The seminar brought participants with diverse computer science backgrounds together to enhance the methodological basis for conducting simulation studies.

Organization and results. Being one day shorter than typical seminars, the seminar started on Tuesday with a short round of introduction and continued with collecting ideas about achievements and challenges of modeling and simulation from the participants on two pinboards (see Figure 6.6). Three talks and partly extensive discussions followed, one focusing on modeling and model-based approaches applied to simulation studies, one on high-performance computing for simulation, and one on analysis and experiment designs. In the late afternoon, the information gathered on the pinboards was revisited. In the end three working groups formed to work towards state-of-the-art and open-challenges papers on the following topics:

- Intelligent Modeling and Simulation Lifecycle
- Policy by simulation: seeing is believing for interactive model co-creation and effective intervention
- Context, composition, automation, and communication: towards sustainable simulation studies

Among the application fields as diverse as cell biological systems, traffic systems, or computer networks, one application dominated the discussions, i.e., Covid-19 simulation. The Covid pandemic showed the importance of modeling and simulation studies being conducted in an efficient, reliable manner, and, accordingly, of comprehensive, intelligent computer support for these studies, it revealed limitations, including those referring to communicating effectively modeling and simulation studies and their results to decision-makers. The results of the working groups are included as short summaries in the full report. Wednesday afternoon, the participants presented their current research work and ideas in a series of lightning talks whose abstracts are also included in the report. However, most of the time was dedicated to the working groups. Plenary sessions on Thursday and finally on Friday allowed the participants to catch up with ideas and the progress made in the different working groups.

	Monday, 10/03	Tuesday, 10/04	Wednesday, 10/05	Thursday, 10/06	Friday, 10/07
08 a.m.		Breakfast	Breakfast	Breakfast	Breakfast
09 a.m.		Introduction incl. identifying challenges	Working groups	Intermediate results	Summary/Work.G.
10 a.m.		30 min. Coffee	30 min. Coffee	30 min. Coffee	30 min. Coffee
11 a.m.		Model-based approaches for Modeling and Simulation	Working groups	Working groups	Final summary
12 noon		Lunch	Lunch	Lunch	Lunch
01 p.m.		Visual analytics, experiment design, statistical model checking	Lightning talks	Working groups	Departure
02 p.m.		10 min. Coffee			
03 p.m.		High Performance Computing: exploiting new architectures			
04 p.m.	Arrival	30 min. Coffee	30 min. Coffee	30 min. Coffee	
05 p.m.		Collecting working group ideas	Hiking	Working groups	
06 p.m.		Dinner	Dinner	Dinner	

Fig. 6.6
Schedule of the Dagstuhl Seminar *Computer Science Methods for Effective and Sustainable Simulation Studies*.

6.59 Foundations for a New Perspective of Understanding Programming

Organizers: André Brechmann, Bonita Sharif, Janet Siegmund, and Westley Weimer

Seminar No. 22402

Date: October 3–7, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.10.61

Participants: Sven Apel, Andrew Begel, Annabelle Bergum, André Brechmann, Teresa Busjahn, Martha E. Crosby, Sarah D'Angelo, Madeline Endres, Sarah Fakhoury, Thomas Fritz, Lena A. Jäger, Timothy Kluthe, Takatomi Kubo, Yun-Fei Liu, Jürgen Mottok, Maria Mercedes T. Rodrigo, Bonita Sharif, Janet Siegmund, Andreas Stefik, Jan Stelovsky, Christine Lourrine S. Tablatin, Westley Weimer, Marvin Wyrich

Remote Participants: Andrew Duchowski, Russell Poldrack

The goal of the seminar *Foundations for a New Perspective of Understanding Programming* was to address how to modernize the perspective on program comprehension and thus make progress regarding our understanding of it. We focused on two challenges: First, we discussed how to provide guidance on addressing methodological issues in interdisciplinary program comprehension research. Second, we aimed at defining a unifying enumeration of the dimensions of a neuroscientific perspective on program comprehension, such that researchers are able to more systematically investigate gaps in the literature.

Through individual participant presentations and the resulting group discussions, we identified several relevant aspects that we discussed further in dedicated working groups. These included discussing how to make better use of eye tracking (Section 4.1 of the full report), identifying the role of readability for program comprehension (Section 4.2 of the full report), and considering how machine learning could help to develop a model of program comprehension (Section 4.3 of the full report). These working groups helped us address the first goal of the seminar by providing guidance on program comprehension research methodologies and identifying potential next steps. To conclude the seminar, participants discussed a possible taxonomy for program comprehension research (Section 5.1 of the full report). This taxonomy identifies commonalities and differences across a broad research area, and addresses our second goal of helping to unify the program comprehension research space; it can serve as a starting point for future researchers to build on to develop an understanding of program comprehension. Also, for researchers entering this field, it is a first glimpse of the complexity of understanding and researching program comprehension.

Beyond identifying research problems in program comprehension research, the many collaborative sessions at this seminar generated numerous potential multi-institutional and interdisciplinary collaborations. We hope that, by making progress on the program comprehension research challenges, we help bring this new research direction one step closer to becoming standard in programming research and disseminating it to a wider audience.

6

6.60 Theory and Practice of SAT and Combinatorial Solving

Organizers: Olaf Beyersdorff, Armin Biere, Vijay Ganesh, and Jakob Nordström

Seminar No. 22411

Date: October 9–14, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.10.84

Participants: Jeremias Berg, Olaf Beyersdorff, Armin Biere, Nikolaj S. Bjørner, Benjamin Böhm, Bart Bogaerts, Jonas Conneryd, Susanna de Rezende, Katalin Fazekas, Mathias Fleury, Vijay Ganesh, Mexi Gioni, Ambros M. Gleixner, Malte Helmert, Marijn J. H. Heule, Matti Järvisalo, Mikoláš Janota, Daniela Kaufmann, Antonina Kolokolova, Laura Kovács, Chunxiao (Ian) Li, Meena Mahajan, Ciaran McCreesh, Kuldeep S. Meel, Jakob Nordström, Andy Oertel, Albert Oliveras, Pavel Pudlák, Torsten Schaub, Andre Schidler, Laurent Simon, Friedrich Slivovsky, Martin Suda, Stefan Szeider, Yong Kiam Tan, Dieter Vandesande, Marc Vinyals, Ryan Williams, Emre Yolcu

This Dagstuhl Seminar gathered leading researchers working on applied and theoretical aspects of the satisfiability (SAT) problem in areas like Boolean satisfiability (SAT) solving and proof complexity and computational complexity theory more broadly, as well as representatives from neighbouring areas such as, e.g., satisfiability modulo theories (SMT) solving, maximum satisfiability (MaxSAT) solving, pseudo-Boolean optimization, constraint programming, and mixed integer linear programming (MIP) on the applied side, and from other areas of computational complexity theory such as exact exponential-time algorithms and parameterized complexity on the theoretical side. This was meant to create an environment conducive to exchange of ideas and techniques between different fields of research. Among the goals of the workshop were to develop a better scientific understanding of real-world efficient computation in general and of the starkly different perspective between the theory and practice of *NP*-hard problems in particular, to explore new approaches for SAT and other challenging combinatorial problems that would have the potential to go beyond the current state of the art, and to stimulate a technology transfer between SAT and other related areas.

This workshop is part of a highly successful series starting at the Banff International Research Station (BIRS) in Canada in 2014, and with follow-up editions held at Schloss Dagstuhl in 2015 (Seminar 15171), at the Fields Institute in Toronto, Canada in 2016, and at the BIRS-affiliated Casa Matemática Oaxaca in Oaxaca, Mexico in 2018. After this fifth edition at Schloss Dagstuhl in October 2022, a sixth edition *"Satisfiability: Theory, Practice, and Beyond"* has already been organized at the Simons Institute for the Theory of Computing at UC Berkeley in April 2023 as part of an eponymous two-month scientific program.

Topic of the Workshop

What served as the point of departure of this workshop is one of the most significant problems in all of mathematics and computer science, namely that of proving logic formulas. This is a problem of immense importance both theoretically and practically. On the one hand, it is believed to be intractable in general, and deciding whether this is so is one of the famous million dollar Clay Millennium Problems (the *P* vs. *NP* problem). On the other hand, today so-called SAT solvers are routinely and successfully used to solve large-scale real-world instances in a wide range of application areas (such as hardware and software verification, electronic design automation, artificial intelligence research, cryptography, bioinformatics, and operations research, just to name a few examples).

During the last two decades there have been dramatic – and surprising – developments in SAT solving technology that have improved real-world performance by many orders of magnitude. However, while modern solvers can often handle formulas with millions of variables, there are also tiny formulas with just a few hundred variables that cause even the very best solvers to stumble. The fundamental question of when SAT solvers perform well or badly, and what underlying properties of the formulas influence performance, remains very poorly understood. Other practical SAT solving issues, such as how to optimize memory management and how to exploit parallelization on modern multi-core architectures, are even less well studied and understood from a theoretical point of view.

Perhaps even more surprisingly, the best SAT solvers today are still based on relatively simple methods from the early 1960s (though the introduction of so-called conflict-driven learning in the 1990s was a very important addition), searching for proofs in the so-called resolution proof system. Although other mathematical methods of reasoning are known that are much stronger than resolution in theory, in particular methods based on

algebra (Gröbner bases) and geometry (cutting planes), attempts to harness the power of such methods have mostly failed to deliver significant improvements in practical performance for SAT solving. And while resolution is a fairly well-understood proof system, even very basic questions about these stronger algebraic and geometric methods remain wide open.

This is an interesting contrast to developments in neighbouring areas such as, e.g., constraint programming and mixed integer programming. There, much more powerful methods of reasoning are successfully used to guide the search, but compared to SAT solving the attempts to employ conflict-driven learning have had much less of an impact. Also, while for SAT solvers it is at least possible to understand some aspects of their performance (by analysing proof systems such as resolution), a corresponding theoretical framework for constraint programming and mixed integer linear programming seems to be mostly missing.

In this workshop, we gathered leading researchers working on SAT and other challenging combinatorial optimization problems in order to stimulate an increased exchange of ideas between theoreticians and practitioners. As discussed above, previous editions of this workshop series at the Banff International Research Station, Schloss Dagstuhl, Fields Institute, and Casa Matemática Oaxaca have had a major impact on the involved communities and have helped to create bridges and stimulate the emergence of a joint research agenda. We are happy to report that the October 2022 workshop at Schloss Dagstuhl fully delivered on the expectation of serving as the valuable next step on this journey. During recent years we have already seen how computational complexity theory can shed light on the power and limitations on current and possible future techniques for SAT and other optimization problems, and and that problems encountered on the applied side can spawn interesting new areas in theoretical research. We see great potential for continued interdisciplinary research at the border between theory and practice in this area, and believe that more vigorous interaction between practitioners and theoreticians could have major long-term impact in both academia and industry.

Goals of the Workshop

A strong case can be made for the importance of increased exchange between the two fields of SAT solving on the one hand and proof complexity (and more broadly computational complexity) on the other. Given how many questions would seem to be of mutual interest, it is striking that the level of interaction had been so low until our workshop series started with the first two meetings in Banff in 2014 and Dagstuhl in 2015. Below, we outline some of the concrete questions that served as motivation for organizing our second Dagstuhl Seminar in this series in 2022, and for broadening the scope from Boolean satisfiability to combinatorial solving and optimization in general. We want to stress that this is far from exhaustive, and we believe that one important outcome of the seminar was to uncover also other questions at the intersection of theoretical and applied research, and of different research areas within combinatorial solving and optimization.

What Makes Formulas Hard or Easy in Practice for Modern SAT Solvers?

The best SAT solvers known today are based on the DPLL procedure, augmented with optimizations such as conflict-driven clause learning (CDCL) and restart strategies. The propositional proof system underlying such algorithms, resolution, is arguably the most well-studied system in all of proof complexity.

Given the progress during the last decade on solving large-scale instances, it is natural to ask what lies behind the spectacular success of CDCL solvers at solving these instances. And given that there are still very small formulas that resist even the most powerful CDCL solvers, a complementary interesting question is if one can determine whether a particular formula is hard or tractable. Somewhat unexpectedly, very little turns out to be known about these questions.

In view of the fundamental nature of the SAT problem, and in view of the wide applicability of modern SAT solvers, this seems like a clear example of a question of great practical importance where the theoretical field of proof complexity could potentially provide useful insights. In particular, one can ask whether one could find theoretical complexity measures for formulas that would capture their practical hardness in some nice and clean way. Besides greatly advancing our theoretical understanding, answering such a question could also have applied impact in the longer term by clarifying the limitations, and potential for further improvements, of modern SAT solvers.

Can Proof Complexity Shed Light on Crucial SAT Solving Issues?

Understanding the hardness of proving formulas in practice is not the only problem for which more applied researchers would welcome contributions from theoretical computer scientists. Examples of some other possible practical questions that would benefit from a deeper theoretical understanding follow below.

- Firstly, we would like to study the question of memory management. One major concern for clause learning algorithms is to determine how many clauses to keep in memory. Also, once the algorithm runs out of the memory currently available, one needs to determine which clauses to throw away. These questions can have huge implications for performance, but are poorly understood.
- In addition to clause learning, the concept of restarts is known to have decisive impact on the performance on modern CDCL solvers. It would be nice to understand theoretically why this is so. The reason why clause learning increases efficiency greatly is clear – without it the solver will only generate so-called tree-like proofs, and tree-like resolution is known to be exponentially weaker than general resolution. However, there is still ample room for improvement of our understanding of the role of restarts and what are good restart strategies.
- Given that modern computers are multi-core architectures, a highly topical question is whether this (rather coarse-grained) parallelization can be used to speed up SAT solving. While there are some highly successful attempts in parallelizing SAT for solving theoretical problems in, e.g., extremal combinatorics, the speed-ups obtained for more applied problems are rather modest or sometimes non-existent. This is a barrier for further adoption of SAT technology already today, and will become a more substantial problem as thousands of cores and cloud computing are becoming the dominant computing platforms in the future. A theoretical understanding of if and how SAT can be parallelized will be essential to develop new parallelization strategies to adapt SAT to this new computing paradigm.

We believe that progress on any of these questions has the potential of influencing the further development of both theoretical and applied research, and to stimulate a further cross-pollination between these two areas.

Can we build SAT Solvers based on Stronger Proof Systems than Resolution?

Although the performance of modern CDCL SAT solvers is impressive, it is nevertheless astonishing, not to say disappointing, that the state-of-the-art solvers are still based on resolution. This method lies close to the bottom in the hierarchy of propositional proof systems, and there are many other proof systems based on different forms of mathematical reasoning that are known to be strictly stronger. Some of these appear to be natural candidates on which to build stronger SAT solvers than those using CDCL.

In particular, proof systems such as polynomial calculus (based on algebraic reasoning) and cutting planes (based on geometry) are known to be exponentially more powerful than resolution. While there has been some work on building SAT solvers on top of these proof systems, progress has been fairly limited. We believe it is fruitful to discuss what the barriers are that stops us from building stronger algebraic or geometric SAT solvers, and what is the potential for future improvements. An important part of this work would seem to be to gain a deeper theoretical understanding of the power and limitations of these proof methods. Here there are a number of fairly long-standing open theoretical questions. At the same time, only in the last couple of years proof complexity has made substantial progress, giving hope that the time is ripe for decisive break-throughs in these areas.

Can Technology Be Transferred Between Different Combinatorial Optimization Paradigms?

Continuing the discussion of stronger methods of reasoning, it is natural to ask whether techniques from e.g., constraint programming (CP) and mixed integer linear programming (MIP) could be imported into SAT solving and vice versa. At a high level, the main loop of combinatorial solvers in all of these paradigms consists of two phases:

- During the *search phase*, the solver makes *decisions* (guesses) about variable assignments and propagates "obvious" consequences until it either finds a solution or discovers a violated constraint (a *conflict*).
- In case of a conflict, during the *backtracking phase* the solver analyses what went wrong and reverses some decision(s) to remove the violation, after which it switches to search again.

For CP and MIP solvers, significant effort is spent during the search phase on making intelligent decisions and deriving (sometimes not so obvious) consequences. In comparison, the backtracking phase is not so sophisticated. For SAT and pseudo-Boolean (PB) solvers it is exactly the other way around. The decisions during the search phase are done quite naively, and since the constraints are relatively simple, propagations cannot be too strong and are fairly easy to detect. Once a conflict is reached, however, an elaborate *conflict analysis* algorithm combines the constraints involved in this conflict to *learn* a new, globally valid constraint that is added to the input formula.

Could it be possible to make more efficient use of conflict analysis in MIP and CP solving? CDCL-style analysis has already been tried with some success, but since this approach boils down to reasoning with clauses it is provably exponentially weaker than what pseudo-Boolean techniques can offer. In the other direction, a quite tempting proposition is to to integrate into SAT and PB solvers the vastly stronger propagation used in CP and MIP. Another change of perspective would be to turn *core-guided MaxSAT solving* into a general pseudo-Boolean optimization technique, using PB conflict analysis to extract better cores, and letting these cores serve as a heuristic for introducing new variables in the spirit of *extended resolution*.

Organization of the Workshop

The scientific program of the seminar consisted of 29 presentations. Among these there were 14 50-minute surveys of different core topics of the workshop. These talks occupied most of the morning schedule Monday-Thursday, and were intended to make sure that the diverse audience would have a bit of a common background for the more technical talks reporting on recent research projects. The list of survey talks and speakers were as follows:

- SAT: interactions between theory and practice (Olaf Beyersdorff)
- SAT and computational complexity theory (Ryan Williams)
- Proof complexity and SAT solving (Jakob Nordström)
- Efficient proof search in proof complexity, a.k.a. automatability (Susanna de Rezende)
- Satisfiability modulo theories (SMT) solving (Nikolaj Bjørner)
- Quantified Boolean formula (QBF) solving and proof complexity (Meena Mahajan)
- Constraint programming (Ciaran McCreesh)
- Mixed integer linear programming (Ambros Gleixner)
- Algebraic methods for circuit verification (Daniela Kaufmann)
- First-order theorem proving (Laura Kovacs and Martin Suda)
- Automated planning (Malte Helmert)
- Formally verified combinatorial solvers (Mathias Fleury)
- Certifying solvers with proof logging (Armin Biere)
- Formally verified proof checking for certifying solvers (Yong Kiam Tan)

The rest of the talks were 25-minute presentations on recent research of the participants. The time after lunch each day was left for self-organized collaborations and discussions, and there was no schedule on Wednesday afternoon.

Based on polling of participants before the seminar week, it was decided to have an open problem session on Thursday afternoon. The poll also asked whether a panel discussion should be organized, but the support for this idea was weaker, and several participants emphasized that the workshop program should not be too dense and that the evenings should be left free of any program. Therefore, the organizers decided not to have a panel discussion. As a nice contribution, some of the participants of the workshop organized a music night on the last evening of the workshop.

6.61 Intelligent Security: Is "AI for Cybersecurity" a Blessing or a Curse

Organizers: Lejla Batina, Annelie Heuser, Nele Mentens, Stjepan Picek, and Ahmad-Reza Sadeghi

Seminar No. 22412

Date: October 9–14, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.10.106

Participants: Ileana Buhan, Lukasz Chmielewski, Alexandra Dmitrienko, Elena Dubrova, Oguzhan Ersoy, Hossein Fereidooni, Fatemeh Ganji, Houman Homayoun, Domagoj Jakobovic, Dirmanto Jap, Florian Kerschbaum, Marina Krcek, Jesus Luna Garcia, Damien Marion, Luca Mariot, Nele Mentens, Irina Nicolae, Stjepan Picek, Jeyavijayan Rajendran, Ahmad-Reza Sadeghi, Patrick Schaumont, Matthias Schunter, Mirjana Stojilović, Shahin Tajik, Trevor Yap

In recent years, artificial intelligence (AI) has become an emerging technology to assess security and privacy. Moreover, we can see that AI does not represent "only" one of the options for tackling security problems but instead a state-of-the-art approach. Besides providing better performance, AI also brings automated solutions that can be faster and easier to deploy but are also resilient to human errors. We can only expect that future AI developments will pose even more unique security challenges that must be addressed across algorithms, architectures, and hardware implementations. While there are many success stories when using AI for security, there are also multiple challenges. AI is commonly used in the black-box setting, making the interpretability or explainability of the results difficult. Furthermore, research on AI and cybersecurity commonly look at the various sub-problems in isolation, mostly relying on best practices in the domain. As a result, we often see how techniques are "reinvented", but also that strong approaches from one application domain are introduced to another only after a long time.

The Dagstuhl Seminar 22412 on 'Intelligent Security: Is "AI for Cybersecurity" a Blessing or a Curse' brought together experts from diverse domains of cybersecurity and artificial intelligence with the goal of facilitating the discussion at different abstraction levels to uncover the links between scaling and the resulting security, with a special emphasis on the hardware perspective. The seminar started with two days of contributed talks by participants. At the end of the second day, every participant suggested topics to be discussed in more detail. From the initial pool of nine topics, we decided to concentrate on four topics on the third and fourth day of the seminar: 1) the explainability of AI for cybersecurity, 2) AI and implementation attacks, 3) AI and fuzzing, and 4) the security of machine learning. The first group approached the problem of the explainability of AI for cybersecurity. The discussion mainly revolved around scenarios where deep learning is used as the attack method, but explainability is necessary to understand why the attack worked and, more importantly, how to propose new defense mechanisms that will be resilient against such AI-based attacks. During the discussion, we considered two perspectives: a) understanding the features and b) understanding deep neural networks.

The second group focused on how AI can improve the performance of implementation attacks. More precisely, we discussed the side-channel analysis and fault injection. Most of the discussion aimed at usages of deep learning for side-channel analysis and evolutionary algorithms for fault injection. However, we also discussed how the lessons learned from one domain could be used in another one. The third group worked on the topic of security fuzzing. We discussed how techniques like evolutionary algorithms are used for evolving diverse mutations and mutation scheduling. At the same time, machine learning is (for now) somewhat less used, but there are many potential scenarios to explore. For instance, instead of using evolutionary algorithms, it should be possible to use reinforcement learning to find mutation scheduling. The fourth group discussed the topic of the security of machine learning. More precisely, it focused on backdoor attacks and federated learning settings. While both attack and defense perspectives were discussed, the discussion group emphasized the need for stronger defenses. Each group followed a cross-disciplinary setting where the participants exchanged groups based on their interests. We had one group switch per day to allow sufficient time to discuss a topic. At the end of each day, all participants joined a meeting to discuss the findings and tweak the topics for the discussion groups. On the last day of the seminar, all participants worked together on fine-tuning the findings and discussing possible collaborations. The reports

of the working groups, gathered in the full report, constitute the main results from the seminar. We consider them the necessary next step toward understanding the interplay between artificial intelligence and cybersecurity, as well as the interplay among diverse cybersecurity domains using AI. Moreover, we expect that the seminar (and the related Dagstuhl Report) will help better understand the main open problems and how to use techniques from different domains to tackle cybersecurity problems. This will encourage innovative research and help to start joint research projects addressing the issues.

6.62 Security of Decentralized Financial Technologies

Organizers: Joseph Bonneau, Arthur Gervais, and Marie Vasek

Seminar No. 22421

Date: October 16–21, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.10.129

Participants: Svetlana Abramova, Sarah Azouvi, Alex Biryukov, Rainer Böhme, Stefanos Chaliasos, George Danezis, Markus Dürmuth, Jens Ernstberger, Bryan Ford, Arthur Gervais, Lioba Heimbach, Philipp Jovanovic, Aljosha Judmayer, Ghassan Karame, Lucianna Kiffer, Ben Livshits, Pedro Moreno-Sanchez, Joachim Neu, Tim Ruffing, Florian Tschorsch, Marie Vasek, Roger Wattenhofer, Aviv Yaish, Fan Zhang, Liyi Zhou, Aviv Zohar

Trusted intermediaries have been the backbone of economic transactions for centuries. However, with the rise of decentralized ledgers like Bitcoin and Ethereum, individuals now have the opportunity to trade and interact without relying on a centralized authority. In 2020, the decentralized finance (DeFi) sector grew to become a 13+ billion USD economy, encompassing exchanges, borrowing/lending, margin trading, derivatives, and more.

The non-custodial nature of decentralized ledgers gives individuals more control over their assets, but it also requires them to take greater responsibility for managing their private keys and assets. Cryptographers expect DeFi users to have a deep understanding of the security properties and guarantees of the protocols, but in reality, it is challenging to keep users informed about these complexities. Therefore, there is a pressing need for more research to clarify user comprehension of DeFi properties. Additionally, decentralized ledgers face a number of technical limitations, such as scalability issues and potential vulnerabilities to pseudonymous malicious actors.

To address these challenges, the Dagstuhl Seminar brought together researchers with expertise in various subfields of DeFi to jointly examine the security, privacy, and financial properties of decentralized finance. The primary objective of the seminar was to explore how to protect DeFi users. The seminar aimed to reconcile the conflicting demands of security, usability, and performance in DeFi, and outline best practices for users to remain safe while engaging in DeFi activities. Finally, the seminar aimed to apply its recommendations to the growing DeFi ecosystem.

During the seminar, participants presented talks on a wide range of topics, including active attacks on the DeFi ecosystem, proposed cryptographic schemes for enhancing the security of cryptocurrencies, and network insights on cryptocurrencies. The seminar also featured productive discussions across working groups, bringing together researchers from diverse perspectives to achieve the common goal of securing the DeFi landscape.

Given the rapid growth of the DeFi sector, it is important to keep exploring ways to improve its security and user-friendliness. One way to do this is through collaboration and information-sharing among researchers, developers, and users. The Dagstuhl Seminar was an important step in this direction, but there is still much work to be done.

One area of focus could be on improving user education and awareness. This could include developing easy-to-understand guides and tutorials, as well as increasing the transparency of DeFi protocols and the risks associated with using them. Additionally, there is a need for more research into the scalability and performance limitations of decentralized ledgers, as well as finding ways to mitigate security risks such as smart contract vulnerabilities.

Another important aspect to consider is the regulatory landscape for DeFi. Currently, many DeFi protocols operate in a regulatory gray area, and it is important to ensure that they comply with relevant laws and regulations while also protecting user privacy and security. This may require more collaboration between DeFi developers and regulators to establish clear guidelines and standards.

Despite the progress made in the DeFi sector, there are still many unknowns that need to be explored. For example, there is limited understanding of how the Ethereum Proof-of-Stake (PoS) security mechanism works, and what guarantees it provides. This is a crucial aspect of the DeFi landscape as Ethereum is the most widely used blockchain for DeFi applications. Further research is needed to understand the security properties of Ethereum PoS and how it can be improved to better protect users. Additionally, there are other areas in DeFi that require further investigation, such

as the scalability and performance limitations of decentralized ledgers, and the trade-offs between privacy and security. By exploring these unknowns, we can gain a better understanding of the DeFi ecosystem and find ways to improve its security and user-friendliness.

In conclusion, the DeFi sector is still in its early stages, and there is much room for growth and improvement. By continuing to bring together experts from various fields and encouraging collaboration, we can help to build a secure and user-friendly DeFi ecosystem that benefits everyone.

6.63 Developmental Machine Learning: From Human Learning to Machines and Back

Organizers: Pierre-Yves Oudeyer, James M. Rehg, Linda B. Smith, and Sho Tsuji

Seminar No. 22422

Date: October 16–21, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.10.143

Participants: Thomas Carta, David J. Crandall, Alejandrina Cristia, Rhodri Cusack, Hana D'Souza, Maureen de Seyssel, Emmanuel Dupoux, Abdellah Fourtassi, Michael C. Frank, Hiromichi Hagihara, Uri Hasson, Felix Hill, Judy Hoffman, Celeste Kidd, Eon-Suk Ko, Maithilee Kunda, Marvin Lavechin, Casey Lew-Williams, Atsushi Nakazawa, Pierre-Yves Oudeyer, Marc'Aurelio Ranzato, James M. Rehg, Clement Romac, Rebecca Saxe, Olivier Sigaud, Stefan Stojanov, Jelena Sucevic, Daniel Swingley, Ngoc Anh Thai, Ingmar Visser, Anne Warlaumont, Gert Westermann, Chen Yu

Remote Participants: Kristen Grauman, Jitendra Malik, Linda B. Smith, Sho Tsuji, Andrew Zisserman

Recent advances in artificial intelligence, enabled by large-scale datasets and simulation environments, have resulted in breakthrough improvements in areas like object and speech recognition, 3D navigation, and machine translation. In spite of these advances, modern artificial learning systems still pale in comparison to the competencies of young human learners. The differences between human learning and the paradigms that currently guide machine learning are striking. For example, children actively identify both the concepts to be learned and the data items used for learning, they infer the labels for learning from ambiguous perceptual data, and they learn from continuous streams of percepts without storing and curating large datasets. Artificial intelligence researchers are increasingly looking to developmental science for ideas and inspiration to improve machine learning, while developmental scientists are adopting tools from data science and machine learning to analyze large datasets and gain insights into developmental processes.

This seminar created new connections between the developmental and machine learning research communities by bringing together researchers in linguistics, psychology, cognitive science and neuroscience with investigators working in computer vision, machine learning and robotics. The seminar focused on three research questions:

1. What are the key computational problems and challenges that need to be addressed in creating a developmentally-inspired machine learner? Existing machine learning methods are built on a set of canonical problem formulations such as supervised learning and reinforcement learning. At the same time, decades of research in developmental science have produced an increasingly detailed characterization of learning in children. How can we leverage these insights to create new and more powerful machine learners and revise standard ML problem formulations?
2. What criteria are necessary for agent-based simulation models of development to advance machine learning and provide useful tests of developmental hypotheses? Advances in computer graphics and physics simulation have made it possible to create synthetic environments for training reinforcement learning agents to perform developmentally-relevant cognitive tasks such as navigating 3D space and manipulating objects. Can such computational experiments serve as useful tests of developmental hypotheses?
3. How can data-driven computational models be used to advance developmental science? It is increasingly feasible to collect dense sensor data that captures the perceptual inputs children receive (e.g., via wearable cameras and eye trackers), their behaviors during naturalistic interactions, and a variety of contextual variables relevant to cognitive tasks. These rich datasets, in conjunction with advances in deep learning have created the opportunity to create machine learning models which can "solve" certain developmental tasks such as object recognition. Given that such deep models do not speak directly to mechanisms of human learning, how can such research advance developmental science?

Through a seminar program consisting of tutorials, talks, working group meetings, and early career mentorship sessions, we gained interdisciplinary insights into these core research questions. Attendees discussed the potential research directions that different research disciplines can benefit from each other, as well as collaboration opportunities and future development of the community. As the initial step, we aim to connect interested researchers online through social media and provide a common repository for relevant literature.

6.64 Data-Driven Combinatorial Optimisation

Organizers: Emma Frejinger, Andrea Lodi, Michele Lombardi, and Neil Yorke-Smith

Seminar No. 22431

Date: October 23–28, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.10.166

Participants: Karen Aardal, Claudia D'Ambrosio, Bistra Dilkina, Ferdinando Fioretto, Emma Frejinger, Maxime Gasse, Stefano Gualandi, Oktay Gunluk, Tias Guns, Serdar Kadioglu, Lars Kotthoff, Hoong Chuin Lau, Pierre Le Bodic, Andrea Lodi, Marco Lübbecke, Sofia Michel, Andrea Micheli, Ruth Misener, Laurent Perron, Sebastian Pokutta, Louis-Martin Rousseau, Helge Spieker, Kevin Tierney, Pashootan Vaezipoor, Pascal Van Hentenryck, Stefan Voß, Neil Yorke-Smith, Yingqian Zhang

In the last five years, an area now being influenced in a new way by machine learning (ML) is combinatorial optimisation (CO). Combinatorial optimisation is studied for both its importance in theory, since CO problems are NP-hard problems, and for its importance in real-world decisions, for example, planning drivers and routes for a fleet of delivery vehicles. CO problems are studied in operations research (OR) and also traditionally in symbolic artificial intelligence (AI) such as constraint programming (CP) and satisfiability modulo theories.

This Dagstuhl Seminar built on the fast-growing interest in combining ML with 'traditional' AI methodologies like CP, and with OR more generally [1, 2]. Surveying the scattered initiatives, the seminar had the ambition to set the agenda for constraint-based 'Combinatorial Optimisation 2.0'. Historically, several communities have focussed on different approaches to CO, mostly in a disjoint manner. This division between, on the one hand, the OR and symbolic AI communities, and on the other, the ML and functional AI communities, is historically strong. While in recent years a dialogue between symbolic and functional AI communities has emerged, there remains too little connection between the discrete OR and ML communities.

The seminar was organised by Emma Frejinger (Canada), Andrea Lodi (USA), Michele Lombardi (Italy) and Neil Yorke-Smith (Netherlands). Michele was unable to attend in person, due to last minute circumstances, and joined plenary parts of the seminar online. Similarly, it was necessary for Pierre-Luc Bacon to give his tutorial remotely.

Seminar Overview

The seminar opened with four tutorials, whose abstracts are given in the full report, on the topics of CP (by Tias Guns), mixed (non)-linear integer programming (MIP) (by Ruth Misener), end-to-end ML for CO (by Ferdinando Fioretto), and reinforcement learning (RL) (by Pierre-Luc Bacon).

The seminar included a set of informal short introductory and topical talks, and sessions of collaborative planning. The overarching questions that structured this planning are, on the one hand, (1) how ML can help in modelling or solving CO problems – or both modelling and solving – and in particular constraint-based models and solving; and on the other hand, (2) how CO can help in tasks approached using ML, including ML training and algorithms. Then, (3) what problems and tasks can be addressed (only) by the synergistic combinations of these methodologies?

Through discussions, the participants identified jointly six topics to be approached in smaller working groups: (i) self-supervised representation learning for combinatorial optimisation, (ii) uncertainty, prediction, optimisation and decision-focussed learning, (iii) OR for ML, (iv) vehicle routing and the role of ML, (v) ML-augmented MIP solvers, and (vi) fairness. The groups discussed challenges, existing work and identified open research questions with promising future avenues at the intersection between OR and ML. The working groups are summarised below.

[47] On Slack, `ML<>CO`

The outcomes of the seminar in furthering the development of a community at the intersection of OR and ML are expected to be felt in the coming couple of years. Already, however, there are tangible outcomes in terms of roadmap ideas, an open online discussion forum (Slack)[47], multiple new collaborations, and a research grant submitted. A special issue of the journal *Frontiers in Applied Mathematics and Statistics* is organised by one of the participants.

The scientific programme was beautifully facilitated by the surroundings and academic services of Schloss Dagstuhl. Further, on the opening evening of the seminar, volunteers among the participants took part in "slide bingo". During this humorous session they improvised presenting the slides of others. On Wednesday afternoon, the participants took a walk to a nearby village in unexpectedly fine sunny weather for October.

Reflections on the Week

All communities present at the seminar found benefit from the interactions and discussions. Meeting in person at the scale of a Dagstuhl Seminar was much appreciated! Participants were aware of the differing emphases, mindsets, and publication practices of different communities. In general, it was felt that strengthening the connection between ML and OR helps in bridging the gap between predictive and prescriptive analytics, which can benefit industrial or government actors and citizens alike.

Working Groups

In this section, we briefly summarise the discussions in the six working groups.

Self-supervised representation learning for combinatorial optimisation

This working group enjoyed lively discussions around the concept of a "foundational model" for CO. The motivation is to avoid retraining from scratch when there is a relatively small change. The group discussed transfer learning in terms of problem formulation, downstream task and instance distributions.

The group identified open questions:

- What is the equivalent of saving models/checkpoints in ML or natural language processing (NLP) in data-driven CO (DDCO)?
- Can we share pre-trained models to generate SAT/CP/MIP embeddings without training again?
- NLP has the concept of a tokeniser that preprocesses the text before it gets fed into the network; in DDCO we would need similar pre processors that transfer the problem instances into the model's expected (graph) structure.
- What is the equivalent of "large" aspect from large language NLP models for DDCO?
- Is there a (super) GLUE benchmark equivalent for DDCO?

Uncertainty, prediction, optimisation and decision-focussed learning

Decision-focussed learning aims at training prediction models against a loss reflecting the quality of decisions instead of a classic prediction loss. This working group weighed up the questions: when is decision-focussed learning (DFL) better than (traditional) alternatives? The group gave energy into thinking about stochastic formulations, data perturbation and interpretability of DFL. The group also identified a connection with RL, in particular contextual bandits (single-stage decision). Since there has been some confusion around the terminology, the group recommended to use "decision-focussed learning" instead of "predict and optimise" or "predict + optimise".

The group wrote down example problems for three settings of decision focussed learning for CO. First, unknown parameters in the objective. This is the most studied case and there are several applications in the literature. Second, unknown parameters in the right-hand side of the constraints. For example, transport network planning where demand predictions occur in capacity constraints. Third, unknown parameters in the left-hand side of the constraints. For example, healthcare scheduling problems where treatment durations are predicted and should not exceed a given schedule length.

Operations research for machine learning

This working group was provoked by the feeling in the OR community that the ML community is seldom happy with discrete optimisation. In other words, how can CO have an influence on problems that generally come from the ML community? Those from OR background in the group expressed that they want to make real contributions to ML.

The group proceeded to outline three major obstacles:

- Scalability. OR methods tend to be limited when dealing with extremely large datasets that are often associated with ML applications.
- Optimality. OR methods have been designed in general to provide guarantees. Computing confidence bounds for ML applications would be excellent but one major obstacle is that the loss function is computed over samples from an unknown distribution. In other words, there is uncertainty with respect to the real objective function, which would require a re-interpretation of what confidence bounds are.
- Software. Whereas the ML community is used to working with self-installing open-source software, the OR community uses more cumbersome, often commercial software.

The recommendation was for research into a new generation of optimisation-based heuristics. The group identified four areas in which discrete optimisation methods are likely relevant: (i) optimal transport problems, (ii) neural-network verification, (iii) the broad area of fairness, explainability and interpretability, and (iv) training for Gaussian processes with tree kernels.

Vehicle routing and the role of machine learning

This working group ascertained exciting research on using ML to help solve routing problems. Two main aspects are, first, that the current state of (deployed) routing software has no idea whether it has seen a problem before, the types of problems being solved, and so forth; and second, anticipating the future – for instance dynamic settings, demand estimation, service time estimation, and so forth – would make the solution of routing problems even more relevant in practice. The working group felt that leveraging ML in both aspects could lead to significant improvements.

The group identified open questions:

- Does the ML model output individual actions or does it output an instance-specific heuristic?
- Can we learn insights about the problem from the predictions?
- Where does or could deep RL work best?
- Can we make a unified routing model (a 'foundation model')?

Machine learning augmented MIP solvers

This working group started from the general questions: which is the big challenge in MIP solving? And, will ML-augmented MIP solvers be ever significantly better than improved versions of the current solvers?

The group found that one significant motivation to go for ML-augmented MIP is 'democratisation': ML could allow the more general use of MIP technology by automatising some steps of MIP development and solution that depend on the specifics of a class of problems in hand without requiring the intervention of experts in the loop. Such a democratisation would require the definition of a robust pipeline on how to learn – from data – tasks like branching, cutting, preprocessing, etc. on the specific class of instances in hand, i.e., characteristic of the application one wants to solve. Such a robust pipeline does not exist yet though there is strong evidence of successful stories where ML-augmented tasks are performed more efficiently than in classical MIP solvers. Some of those successful examples have been already integrated into the solvers, even commercial ones.

The group identified a number of interesting directions, as yet unexplored in the fast-moving subfield of ML-for-MIP. Among these are hypothesis generation, defining appropriate performance metrics, learning for cutting plane generation and selection. The group also discussed benchmark libraries.

Fairness

This working group sought to learn more about data-driven CO models for fairness. The group recognised that an issue with fairness is already in its definition. While in the ML community, the concept of fairness is related to a tradeoff between overall accuracy and group accuracy, in CO for decision making, there is not a clear definition of fairness.

The group discussed an application in the online scheduling of radiologists and neuro-radiologists of CT scans. In this context, because of the scarcity of the resources, fairness is associated with their correct and 'fair' use.

Fairness at large is also related to explainability and interpretability and the working group discussed the use of classical CO methods that tend to be more interpretable of the ML ones that are often perceived as black-boxes. Further, a potential important area in this context is that of integrating ML and OR to achieve a higher level of explainability, for example by improving methods like decision trees and ML classification algorithms.

References

1 Yoshua Bengio, Andrea Lodi, Antoine Prouvost: *Machine learning for combinatorial optimization: A methodological tour d'horizon*. Eur. J. Oper. Res. 290(2): 405–421 (2021). https://doi.org/10.1016/j.ejor.2020.07.063

2 James Kotary, Ferdinando Fioretto, Pascal Van Hentenryck, Bryan Wilder: *End-to-End Constrained Optimization Learning: A Survey*. IJCAI 2021: 4475–4482. https://doi.org/10.24963/ijcai.2021/610

6.65 Towards a Unified Model of Scholarly Argumentation

Organizers: Khalid Al-Khatib, Anita de Waard, Iryna Gurevych, and Yufang Hou

Seminar No. 22432

Date: October 23–28, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.10.175

Participants: Khalid Al-Khatib, Milad Alshomary, Wolf-Tilo Balke, Tilman Beck, Elena Cabrio, Fengyu Cai, Davide Ceolin, Anita de Waard, Nils Dycke, Dayne Freitag, Daniel Garijo, Iryna Gurevych, Graeme Hirst, Yufang Hou, Eduard H. Hovy, Anne Lauscher, Maria Liakata, Tobias Mayer, Robert Mercer, Smaranda Muresan, Preslav Nakov, Sukannya Purkayastha, Chris Reed, Domenic Rosati, Florian Ruosch, Harrisen Scells, Ferdinand Schlatt, Benno Stein, Simone Teufel, Serena Villata, Andreas Vlachos, Henning Wachsmuth, Ryan Wang

Background

Argumentation is prevalent in scientific discourse and critical to scientific progress. Recent efforts have attempted to identify and model argumentative structures in scholarly discourse from different perspectives. Within the domain of scientific literature analysis, computational approaches to argumentation have followed the route of discourse modeling by identifying relations between spans and clauses encoding rhetorical structures (e.g., premises and conclusions), or as typed turns in community debate (e.g., supports or attacks). Another thread of research, often applied to biomedical literature, focuses on capturing functional discourse at different levels of granularity, such as objectives, methods, results or scientific claims, and their relation to reported evidence. Most work adopts a corpus perspective, either highlighting the role of sentences or phrases within the scientific discourse or aligning claims across documents, and using citations to construct claim-evidence networks that summarize the state of knowledge in a field. Within the health sciences, argumentative structures have been used to automate the production of systematic reviews by identifying key actionable knowledge elements from collections of clinical reviews, case studies, and research papers. For an overview of previous work, see e.g., [1].

Despite these varied efforts and the clear practical importance of the work, there is lack of consensus on how scientific argumentation should be formalized. For instance, it remains unclear whether formalisms popular in non-scientific domains apply to scientific discourse, and whether a single formalism can adequately support argumentation research in diverse disciplines such as biology, chemistry, materials science, medical research and computer science. This lack of consensus manifests in a dearth of shared reference corpora, which are needed to advance research into computational treatments of scientific argumentation. It has also led to the absence of an operational theory for defining argumentative components in scholarly text.

Goals

Our Dagstuhl Seminar, titled *Towards a Unified Model for Scholarly Argumentation*, sought to further the emergence of this missing consensus. Specifically, the seminar objectives included:

- Enabling robust advances in argument technology by collecting and working on use cases in scholarly and medical discourse;
- Starting the development of a foundational model for argumentation in science and healthcare;
- Laying the groundwork for a multidisciplinary community devoted to building and maintaining principles, tools, and models to identify key components in scholarly argumentation.

Outcomes

The seminar was attended by scientists at different levels of seniority and from a variety of research backgrounds. Some participants have made the computational modeling of argumentation or the scholarly literature the central focus of their careers. Others were drawn to the seminar through their work on applications in adjacent problem areas. Ultimately, all emerged with a sense that important bonds of shared interest had formed, fostered by several seminar outcomes.

Knowledge Baselining

A shared understanding of the problem space was obtained, through a series of keynotes and panel discussions on theory, models, tools, and available corpora. These are described in greater detail in the full report, in Section 3. In particular, two introductory talks summarized the state of the art in argument modeling (3.1) and computational argument mining (3.2).

Five further plenary talks described different use cases where argument identification can support NLP tasks:

- using scientific discourse to understand and measure the impact of scholarly contributions (3.3);
- using argument modeling to generate discourse (3.4);
- generating scholarly documents using argument structures (3.5);
- interpreting a fortiori arguments (3.6);
- synthesizing evidence from text to support public policy (3.7).

A series of eleven flash talks covered a host of other efforts, presenting corpora, tools, and relevant applications, such as document understanding, extracting high-level claims, and identifying fallacious and persuasive elements in scholarly texts (Section 4).

Problem Elucidation

At the beginning of the workshop, the group identified several important focus areas that then became the subject of breakout group deliberation over the course of the week. All materials, including the full program, slides, summaries of the breakout sessions and code and corpora submitted can be found on the workshop Google Drive at https://bit.ly/TUMSA22.

- *Foundations* (Section 5.1). A subgroup of participants discussed a shared argumentation model, based on the various proposals presented during the plenary sessions. The group debated and wrote a first-order consensus of these varying views, which can be used for further development of a foundational model of scholarly argumentation.
- *Domains* (Section 5.2). This working group pursued a comparison of argumentation in different scholarly domains. A methodology was delineated for how to annotate argumentation across domains while reducing the need for domain experts.
- *Argument Quality* (Section 5.3). This working group explored how argumentation quality can be evaluated, and defined a series of questions to assess this. Additionally, the group members contributed an open-source tool to perform the evaluation of argumentation quality, which can be further developed to support this task.
- *Community Dialogue* (Section 5.4). This working group looked at how argument structure can support an important editorial task, namely to decide on an accept or reject decision for a submitted manuscript, based on a number of peer reviews. The group developed a corpus of dialogues that simulate how a meta-reviewer asks questions about a document that has received a number of reviews, which can be used in future work in this domain.

Community Formation

Building on the connections developed during the seminar, a series of collaborations have been fostered, and thoughts on how to proceed with this work through a multidisciplinary lens have been put forth. Multiple new collaborations have been formed as an outcome of this week, in some cases centered on new tools and research corpora first conceived in the workshop.

Next Steps

This Dagstuhl Seminar brought together a multi-disciplinary, international, and diverse community of researchers from academia and industry to discuss scholarly argumentation. Much argumentation occurred, during and after presentations, in breakout groups, during the social events spread out through the week, and long into the night. Necessarily, this is only the beginning of a conversation that will unfold over the coming years, one that will ultimately produce a shared model of scholarly argumentation and a set of concrete research tasks and important new use cases.

We hope that this seminar was the first in a series of events devoted to this topic, that this inaugural event proves pivotal in the formation of a cohesive research community addressing a problem with large practical ramifications. The Dagstuhl Report of this seminar can hopefully contribute to accelerate work in this area, by offering a summary of current efforts, and a number of interesting problems to work on.

References

1 Khalid Al Khatib, Tirthankar Ghosal, Yufang Hou, Anita de Waard, and Dayne Freitag. 2021. Argument Mining for Scholarly Document Processing: Taking Stock and Looking Ahead. In Proceedings of the Second Workshop on Scholarly Document Processing, pages 56–65, Online. Association for Computational Linguistics.

6.66 Optimization at the Second Level

Organizers: Luce Brotcorne, Christoph Buchheim, Dick den Hertog, and Gerhard J. Woeginger

Seminar No. 22441

Date: October 30– November 4, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.10.207

Participants: Yasmine Beck, Luce Brotcorne, Christoph Buchheim, Martina Cerulli, Claudia D'Ambrosio, Danique de Moor, Dick den Hertog, Boris Detienne, Gabriele Dragotto, Marc Goerigk, Dorothee Henke, Felix Hommelsheim, Quentin Jacquet, Jannis Kurtz, Martine Labbé, Christina Liepold, Frauke Liers, Ivana Ljubic, Ahmadreza Marandi, Komal Muluk, Bernardo Pagnoncelli, Jean Pauphilet, Michael Poss, Krzysztof Postek, Ted Ralphs, Martin Schmidt, Juan Pablo Sepulveda Adriazola, Shimrit Shtern, Nathalia Wolf, Pawel Zielinski

Topic of the Seminar The second level of the polynomial hierarchy contains a variety of problems that allow natural simple formulations with one existential and one universal quantifier. For instance, a typical problem in robust optimization asks whether there EXISTS some production plan that performs reasonably well under ALL possible price scenarios for electricity in the coming two years. A typical problem in bilevel optimization asks whether there EXISTS a way of setting taxes so that ALL possible behaviors of the citizens generate a reasonable tax revenue. A typical problem in Stackelberg games asks whether there EXISTS a starting move for the first player that wins the game against ALL possible counter-moves of the second player.

Problems of this type are usually complete for the class Σ_2^p and hence are most likely not contained in the class NP. For that reason, the methodologies that have been developed for NP-complete problems over the last 50 years do not directly apply to robust and/or bilevel optimization problems. Up to the current moment, most of the work on such problems is purely computational and without any deeper theoretical understanding. Most approaches simply try to carry over the well-developed machinery from integer programming to concrete robust problems and bilevel problems. We will need to develop new techniques, new tricks, new insights, new algorithms, and new theorems to get a grip on this area.

The goal of this Dagstuhl Seminar was to bring together experts in theoretical computer science and experts in combinatorial optimization, and to work towards the following goals:

- summarize the status quo of robust optimization and bilevel optimization,
- identify central research lines on the computational and implementational front,
- identify central research lines in theoretical computer science, as for instance in parameterized complexity and approximability.

The list of participants perfectly reflected these goals, it included experts in complexity theory as well as researchers interested in the development of effective algorithms and the practical solution of real-world bilevel or robust optimization problems.

Implementation and Conclusions In this seminar we brought together, for the first time, leading researchers from three different communities (robust optimization, stochastic optimization, and bilevel optimization) in order to bridge the gap between these fields from a theoretical and practical point of view.

Considering the different backgrounds of participants, we scheduled several talks with an introductory character in the first half of the week: Marc Goerigk and Frauke Liers gave overview talks on robust optimization, from a combinatorial and continuous perspective, respectively. Martine Labbé presented the state of the art of bilevel optimization, while Martin Schmidt combined both topics in his talk on bilevel optimization under uncertainty. Bernardo Pagnoncelli gave an introduction into the related topic of stochastic optimization. These presentations laid a common foundation for all further presentations and discussions and were thus a crucial prerequisite for the success of the seminar.

The contributed talks covered a wide range of topics, including complexity theoretic results for (certain or uncertain) bilevel optimization, new models and new methods for bilevel or robust optimization, and new approaches for solving bilevel or robust optimization problems arising in practice. Apart from the exciting contents of these talks, a particularly positive aspect was the large representation of young researchers. In fact, seven out of 18 contributed talks were given by PhD students.

The whole seminar was marked by a very open and constructive atmosphere and by an extraordinarily interactive approach: many presentations quickly turned into lively discussions involving many different participants, often making the original schedule obsolete, but with the benefit of a better common under-

standing and often new insights. One of the recurrent topics arising in many of these discussions was the connection between bilevel and robust optimization, the two main subjects of the seminar. The fact that both problem classes lead to potentially Σ_2^p-hard problems, as mentioned above, yields a connection of rather theoretical nature. From a more concrete point of view, the discussion was about whether one of the two problem classes can be seen (or modeled) as a special case of the other. Additionally, interesting links to game theory and stochastic optimization were identified.

Even though a conclusive answer to these questions could not be given (and probably does not exist), one of the main insights of the seminar was that bilevel and robust optimization, though being investigated in separate communities, share many structural and algorithmic properties. It is worth studying these connections and sharing the knowledge of both communities in order to profit from one another. The Dagstuhl Seminar on "Optimization at the Second Level" was a first significant step into this direction, which is hopefully followed by further progress.

The seminar was a big success, it stimulated new and very fruitful collaborations. We got laudatory feedback from many participants who were already thinking of organizing another seminar on a the same topic in the future.

In memoriam At this point, our thoughts go out to our late colleague and friend Gerhard J. Woeginger. It was his idea to organize a seminar on "Optimization at the Second Level", and without him the seminar would never have become real. Unfortunately, he could no longer witness how his idea was put into practice and how fruitful it turned out to be.

References

1 Gerhard J. Woeginger: "The trouble with the second quantifier", 4OR – A Quarterly Journal of Operations Research, Vol. 19(2), pp. 157–181, 2021.

6.67 Toward Scientific Evidence Standards in Empirical Computer Science

Organizers: Brett A. Becker, Christopher D. Hundhausen, Ciera Jaspan, Andreas Stefik, and Thomas Zimmerman

Seminar No. 22442

Date: October 30–November 4, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.10.225

Participants: Brett A. Becker, Andrew Begel, Michelle Craig, Andrew Duchowski, Neil Ernst, Arto Hellas, Christopher D. Hundhausen, Ciera Jaspan, Timothy Kluthe, Juho Leinonen, Joseph Maguire, Monica McGill, Brad Myers, Andrew Petersen, Mauro Pezzè, Paul Ralph, Kate Sanders, Andreas Stefik, Claudia Szabo, Jan Vahrenhold, Titus Winters, Aman Yadav

The goals of the seminar *Toward Scientific Evidence Standards in Empirical Computer Science* were to establish a process for introducing evidence standards in computer science, build a community of scholars that discuss what a general standard would include and have enough diversity of background to have a good basis for the breadth of community needs across a range of computer science-related venues.

Over the first few days, we conducted a series of breakout groups and larger group discussions. In these, to introduce people to evidence standards, we reviewed several, including: APA JARS [1], WWC [2], and CONSORT [3]. The purpose was introductory and to scaffold for discussions on what could work across the breadth of computer science or in subareas. We also conducted a session looking at existing papers and noted the changes that would need to be made to fit the APA JARS standards. This exercise in particular was found to be useful by participants, as it made it clear that the the conversion is not particularly difficult, although it is aided by advanced planning for what might need to be collected during a study.

During the Dagstuhl, we also had several talks. These included an introductory talk by Andreas Stefik on evidence standards as a whole, telling the story of the well-known Tolbutamide drug and its influence on the medical field in regard to evidence standards. Christopher Hundhausen provided a talk on his experience with introducing reporting standards at ACM's Transactions on Computing Education (TOCE) (Section 3.2 of the full report). Paul Ralph presented on the problems in scholarly peer review and how evidence standards could be a solution, along with a reviewing tool that he has developed (Section 3.3 of the full report). Neil Ernst covered registered reports, their benefits to the transparency and quality of research, and his experience with introducing them at Mining Software Repositories (MSR) and Empirical Software Engineering (EMSE) (Section 3.4 of the full report). Lastly, Kate Sanders et al. discussed a review on reviews, which spanned a variety of the computer science subfields. This included their observations on review criteria, ethical concerns in the peer review process and excerpts from interviews with conference chairs and journal editors that were relevant to the subject of the seminar (Section 3.5 of the full report). Each of these gave insights into the process of adopting an evidence standard and some of the potential impacts of the status quo and potential changes (positive or negative).

Finally, after discussion, we identified four topics for breakout groups to brainstorm potential avenues toward actionable progress on goals: a deeper dive into how to write guidelines for more complex experiments like mixed-methods studies (Section 4.5 of the full report), how can we measure the effects that evidence standards have both in reference in paper quality and community satisfaction (Section 4.6 of the full report), what are the first steps towards community engagement as far as introducing the topic and adopting it (Section 4.7 of the full report) and how to operationalize these standards in a way that is open source to allow for community control (Section 4.8 of the full report). A final working group session went through some of the first steps could be made at conferences and a dissemination plan for how to start information the community about the topic (Section 4.9 of the full report).

Overall, the seminar brought a range of computer science stakeholders up to speed on the state of evidence standards in the field, what could be gained by moving towards a domain-wide guidelines and started a discussion on how to spark the conversation in various communities. A set of next steps on where and what to recommend and talk about with communities were set in motion, as well as plans for a collaborative position paper to introduce the topic to a wider audience.

References

1 American Psychological Association. APA Style Journal Article Reporting Standards (APA Style JARS). Accessed on December 12, 2022 from https://apastyle.apa.org/jars.

2 National Center for Education Evaluation and Regional Assistance. WWC | Find What Works. Accessed on December 12 , 2022 from https://ies.ed.gov/ncee/wwc/.

3 The CONSORT Group. CONSORT Transparent Reporting of Trials. Accessed on December 12, 2022 from https://www.consort-statement.org/.

6.68 Principles of Contract Languages

Organizers: Dilian Gurov, Reiner Hähnle, Marieke Huisman, and Giles Reger

Seminar No. 22451

Date: November 6–11, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.11.1

Participants: Wolfgang Ahrendt, Bernhard Beckert, Alessandro Cimatti, David Cok, Claire Dross, Gidon Ernst, Dilian Gurov, Reiner Hähnle, Sylvain Hallé, Paula Herber, Asmae Heydari Tabar, Marieke Huisman, Marie-Christine Jakobs, Eduard Kamburjan, Nikolai Kosmatov, Srdan Krstic, Sophie Lathouwers, Martin Leucker, Christian Lidström, Rosemary Monahan, Doron A. Peled, Giles Reger, Kristin Yvonne Rozier, Philipp Rümmer, Thomas Santen, Marco Scaletta, Ina Schaefer, Julien Signoles, Alexander J. Summers, Mattias Ulbrich, Petra van den Bos

The full report documents the program and the outcomes of Dagstuhl Seminar 22451 "Principles of Contract Languages".

Formal, precise analysis of non-trivial software is a task that necessarily must be decomposed. The arguably most important composition principle in programming is the procedure (function, method, routine) call. For this reason, it is natural to decompose the analysis of a program along its call structure. Decomposition in this context means to replace a procedure call with a declarative description, possibly an approximation, of the call's effect. In his seminal work on runtime verification in Eiffel, Bertrand Meyer suggested to use the metaphor of a contract between the user (caller) and implementor (callee) for such a description.

Contracts continue to be a central element in run-time (dynamic) analysis. In the last two decades they also became the dominant decomposition approach in deductive verification and are realized in all major software verification systems. More recently, software contracts are increasingly used in test case generation and model checking. Furthermore, programming languages such as "Racket" or "Dafny" were designed with a notion of contract. Contract-based specification languages are available for mainstream programming languages, notably JML for "Java" and ACSL/ACSL++ for "C"/"C++".

However, there is considerable fragmentation concerning terminology, basic principles, expressivity, and usage of contracts in different research communities. Therefore, this Dagstuhl Seminar convened researchers working with contracts in static verification, runtime verification, as well as testing, with the goal of creating a unified view on software contracts.

The seminar participants discussed the following topics and questions:

1. *Sub-procedural* contracts: contracts for blocks, loops, suspension points, barriers
2. Combining *trace-based specifications* for *global* properties with *two-state* contracts for recursive procedures
3. Rethink *abstract* versus implementation-layer specifications
4. Contracts for *concurrent languages*
5. Contract *composition*
6. Unify contracts for deductive and runtime verification and testing
7. (Behavioral) *types* as lightweight contracts
8. Where do contracts come from? Contract *synthesis*
9. Contract *validation*, connection to *natural language*
10. Contracts for *refinement* (correctness by construction)
11. Contracts for relational properties
12. Debugging contracts
13. Domain-specific contract languages

6

6.69 Computational Social Dynamics

Organizers: Martin Hoefer, Sigal Oren, and Roger Wattenhofer

Seminar No. 22452

Date: November 6–11, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.11.28

Participants: Yakov Babichenko, Gerdus Benadè, Ioannis Caragiannis, Giorgos Christodoulou, Andrei Constantinescu, Michal Feldman, Tobias Harks, Martin Hoefer, Max Klimm, Ron Lavi, Pascal Lenzner, Kevin Leyton-Brown, Vangelis Markakis, Darya Melnyk, Noam Nisan, Sigal Oren, Ori Plonsky, Maria Polukarov, Nidhi Rathi, Rebecca Reiffenhäuser, Jörg Rothe, Daniel Schmand, Giovanna Varricchio, Carmine Ventre, Roger Wattenhofer, Lisa Wilhelmi, Jens Witkowski, Yair Zick

Algorithmic techniques provide a powerful toolbox for understanding many phenomena arising in modern society. Often, these phenomena are related to dynamics (e.g., dynamic information spreading, or dynamics in social networks that result from agent interaction). A large part of the present social interaction on networks can be expressed using game-theoretic or microeconomic concepts, e.g., the dynamics of opinions in networks, pricing problems and viral marketing, network-based effects of opinions, group formation, cognitive bias, etc. These problems are rigorously analyzed in the area of algorithmic game theory, where researchers apply the algorithmic toolbox to analyze various social systems.

In this field, a number of applications have not received sufficient attention, which are recently becoming increasingly prominent. For example,

- issues of fairness and bias are central challenges in modern societies.
- behavioral economists are challenging standard assumptions that humans are maximizing expected utility. This change in perspective also poses new challenges for suitable models and algorithm design.
- an important trend in the modern economy are informational challenges, e.g., in problems involving recommendation, persuasion, delegation, or (smart) contract design.

The main aim of this seminar was to bring together a leading set of researchers to discuss these and other challenges, and to advance the state of the art in several new directions. The majority of participants were academics from computer science departments; some were from other disciplines such as economics, mathematics, or electrical engineering. All participants had strong interdisciplinary interests that typically span economics, game theory, and theoretical computer science.

The seminar started on Monday with an introductory session, in which participants introduced their name, affiliation, main research interest as well as a "crazy idea" or "provoking thought". This session was very well-received by the participants, and it initiated discussions directly from the start. The subsequent program included four invited talks/tutorials of roughly one hour each, on several issues chosen by the organizers.

Monday Caragiannis talked about notions of fairness for resource allocation problems.

Tuesday Feldman presented recent innovations in algorithms for contract design.

Wednesday Plonsky discussed behavioral experiments and ideas to predict human behavior.

Thursday Wattenhofer surveyed issues in e-government, blockchains, and finance.

The contributed talks were solicited from the participants and lasted around 20–25 minutes each. In many of these talks there were lively discussions. They continued in the times from lunch to afternoon coffee, which were free for research and individual collaborative meetings.

There was a substantial set of presenters who focused on aspects of fair division. Rathi discussed epistemic notions of envy-freenes and their efficient computation. Bernadé presented results on online algorithms for fair division and guaranteeing notions of envy-freeness and Pareto-efficiency. Zick showed existence and efficient computation for several fairness criteria for matroid-rank valuations. Varricchio discussed algorithmic aspects of randomization and entitlements in fairness notions. Reiffenhäuser presented results on the quality of equilibria in fair division when strategic misreporting of preferences is allowed.

Another focus was the analysis of dynamics and stability in networks. Ventre discussed the clearing problem in financial networks with credit default swaps and showed FIXP-completeness

results. Wilhelmi presented a game-theoretic model for clearing and results on equilibrium computation. Schmand considered a game-theoretic model for opinion formation in networks and convergence of natural best-response dynamics. Lenzner surveyed recent work on network creation games as well as network games that model segregation aspects.

Further talks addressed a number of different areas. Witkowski studied incentive-compatible forecasting competitions. Lavi considered incentives when several of these contests are available to participants. Klimm surveyed work on impartial selection, e.g., when agents have to decide on a representative member among themselves. Leyton-Brown posed new directions and open problems in the understanding of behavior in mobile gaming. Babichenko discussed aggregation mechanisms for anonymous information. Markakis presented a novel algorithm to compute an equilibrium in 2-player zero-sum games. Melnyk explained issues arising from Byzantine behavior in social choice. Last, but not least, Christodoulou gave a proof of the 23-year old Nisan-Ronen conjecture that every truthful mechanism for unrelated machine scheduling can only guarantee a linear approximation ratio.

The seminar was a big success. We believe it will stimulate new and very fruitful collaborations. We got laudatory feedback from many participants which is also reflected in the survey conducted by Dagstuhl.

We thank Giovanna Varricchio for serving as collector of abstracts and open problems.

6

6.70 Dynamic Graph Algorithms

Organizers: Aaron Bernstein, Shiri Chechik, Sebastian Forster, and Tsvi Kopelowitz

Seminar No. 22461

Date: November 13–18, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.11.45

Participants: Amir Abboud, Aaron Bernstein, Sayan Bhattacharya, Joakim Blikstad, Karl Bringmann, Keren Censor-Hillel, Shiri Chechik, Keerti Choudhary, Aleksander Christiansen, Jeremy Fineman, Nick Fischer, Sebastian Forster, Pawel Gawrychowski, Gramoz Goranci, Fabrizio Grandoni, Kathrin Hanauer, Adam Karczmarz, Peter Kiss, Tsvi Kopelowitz, William Kuszmaul, Jakub Lacki, Nicole Megow, Matthias Mnich, Shay Mozes, Danupon Nanongkai, Yasamin Nazari, Nikos Parotsidis, Richard Peng, Maximilian Probst Gutenberg, Eva Rotenberg, Piotr Sankowski, Thatchaphol Saranurak, Christian Schulz, Chris Schwiegelshohn, Shay Solomon, Clifford Stein, David Tench, Jan van den Brand, Virginia Vassilevska Williams, David Wajc, Oren Weimann, Nicole Wein, Uri Zwick

The field of dynamic graph algorithms has evolved rapidly over the past decade. New techniques, new problems, new lower bounds, and new approaches have yielded an extremely fruitful research environment. This seminar provided a venue for the community to establish the main challenges that remain and to actively shape the direction of the field going forward.

The seminar brought together the leading researchers and "rising stars" of the field as well as experts in "neighboring" areas such as distributed algorithms, parallel algorithms, streaming algorithms, online algorithms, approximation algorithms, data structures, fine-grained and parameterized complexity, and optimization. Many participants were also actively researching algorithms engineering for dynamic graph problems, which added interesting perspectives on the prevalent theory-practice gap and fundamental methodological challenges.

Several participants gave talks that were highlighting "cutting-edge" advances in the field, including results that very recently appeared in top theory venues such as STOC, FOCS and SODA (and in some cases won the best paper award). Some participants explored connections to other related areas in algorithms research such as distributed and streaming algorithms or differential privacy. Many of the talks also included suggestions on future directions and highlighted the main challenges in the area.

In two open problem sessions, the participants explicitly identified several central open problems that continue to resist progress. We hope that the resulting list of open problems will be a valuable resource for future research in this field.

During and after the sessions, attendees participated in vibrant discussions. Such interactions enhanced the overall experience and made a clear distinction to a traditional conference-like format. In addition to the research activities, there were many social activities (such as board games, poker, music night, hiking, and ping pong) that made the workshop also a great networking opportunity, in particular for the relatively large fraction of participants who were first-time visitors to Schloss Dagstuhl.

Acknowledgments The organizers would like to thank Yasamin Nazari and Nicole Wein for helping edit the Dagstuhl Report of this seminar, Monika Henzinger for helping develop the concept of this seminar, and the Dagstuhl team for their first-rate support.

6.71 Set Visualization and Uncertainty

Organizers: Susanne Bleisch, Steven Chaplick, Jan-Henrik Haunert, Eva Mayr, and Marc van Kreveld

Seminar No. 22462

Date: November 13–18, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.11.66

Participants: Daniel Archambault, Michael Behrisch, Susanne Bleisch, Annika Bonerath, Steven Chaplick, Sara Irina Fabrikant, Amy Griffin, Jan-Henrik Haunert, Stephen G. Kobourov, Martin Krzywinski, Eva Mayr, Wouter Meulemans, Silvia Miksch, Martin Nöllenburg, Helen C. Purchase, Peter Rodgers, Christian Tominski, Nathan Van Beusekom, Marc van Kreveld, Markus Wallinger, Bei Wang Phillips, Alexander Wolff, Hsiang-Yun Wu

Research Area

The topic of *Set Visualization and Uncertainty* is inherently interdisciplinary, combining aspects of several diverse fields. As such, the overview of the research area is split into the key fields associated with it; namely, information visualization, set systems, graph drawing, uncertainty (as applied to data sets), and cartography.

Information visualization (InfoVis) can help humans gain insight from large volumes of data by providing good graphical overviews as well as appropriate interfaces for accessing details (see, e.g., [1]). It has thus become of high relevance for industry and many scientific disciplines. Since the generation of effective visualizations requires knowledge of human cognition, algorithms, data characteristics, visual variables, and tasks, the InfoVis community embraces members of various disciplines, including computer scientists of different areas, cognitive scientists, psychologists, and cartographers.

Sets are mathematically defined as unordered collections of distinct objects. They play an important role in InfoVis since reasoning based on aggregated information (i.e., sets instead of individual objects) can greatly reduce the complexity of data analysis tasks. Most often, the sets are defined by categories of objects; e.g., people can be grouped by country of residence, education, or gender to study influences on income. Often, the aim is to visualize statistics (e.g., number of elements, average income) for each set and, since an element can be member of multiple sets, the relationships between them (e.g., intersection and containment). Set visualization is traditionally done with Venn or Euler diagrams, yet a plethora of alternative visualization types for sets has been developed. A recent focus of research has been on developing scalable solutions (e.g., to create effective visualizations for very large set systems) and dealing with dynamics (e.g., changes of the elements' set memberships over time). In this seminar we dealt with a different issue, already relevant for static and small set systems: uncertainty. Although the importance of uncertainty visualization has been stressed by several researchers, only few studies exist that deal with it specifically in the context of sets and systems of sets [3].

Uncertainty is inherent to almost any information collected through observations by humans or sensors. Since the assignment of elements to categories corresponding to sets follows observations, the set memberships are uncertain, too. Moreover, subsuming multiple elements with their individual properties under one category results in a loss of information. Although this information reduction may be intended to reduce the graphical complexity, visualizing the within-set as well as the between-set variability may improve the interpretation of the data. Uncertainty is usually evaluated with statistical methods or concepts of probability. Uncertainty can relate to the existence of an element, the existence of a set, the presence of an element in a set, set containment in hierarchies, location of an object in geo-located data, etc. Moreover, uncertainty can be given as a binary property or as a probability. Fuzzy set theory extends the idea of sets by allowing partial set membership, indicated by a value between 0 and 1. This model has been proposed for concepts that lack crisp boundaries (e.g., "young" and "old" as categories of people). In InfoVis, an important question is whether and how the uncertainty of the information displayed should be visually encoded (e.g., with glyphs or graphical variables), and how users process this visualization of uncertainty [2]. Moreover, although the uncertainty may not be depicted, it may be considered when generating a visualization (e.g., by filtering information based on its certainty). Conversely, a standard visualization like a heat map suggests uncertainty which may not exist in the data at all.

Graph drawing is a branch of computer science focusing on the computation of geometric layouts of graphs, involving both formal and experimental methods. Since graphs are useful mathematical models for networks, graph drawing is of high rele-

vance for network visualization. Graph drawing can be applied to non-geometric networks (e.g., social networks consisting of friendship relationships) as well as to geometric networks (e.g., networks of metro lines) if the aim is to generate more abstract (e.g., schematic) representations. Since a system of sets can be considered as a hypergraph in which each node corresponds to one entity and each hyperedge corresponds to one set, set visualization is fundamental in the graph drawing community. However, aspects of uncertainty remain mostly unexplored [4].

Geographic information is a combination of geometric, temporal, and attribute information, each of which can be uncertain in different ways and can be visualized in different ways commonly through maps. *Cartography* and its sister discipline *Geographic Information Science* have a long history in dealing with uncertainty in the context of analyzing and visualizing spatial information. For example, international standards formalizing elements of spatial data quality have been established (e.g., ISO 19157:2013 defining thematic accuracy, temporal quality, positional accuracy) and graphical variables encoding the uncertainty of information in maps have been proposed, including color saturation and symbol focus [5].

Seminar Goals

This seminar aimed to advance research into methods and techniques for set visualizations and uncertainty by fostering interdisciplinary and cross-domain collaboration (cf. Section "Research Area" in the full report). Sets are mathematically defined as collections of distinct objects. They play an important role in Information Visualization since reasoning based on aggregated information can reduce the complexity of the analysis tasks. Uncertainty is inherent to almost any information collected through observations by humans or sensors and, thus, also set elements or their set membership. Uncertainty generally adds to the complexity of data analysis and data presentation. In the seminar we looked specifically into approaches for dealing with uncertain information when visualizing sets. Information Visualization has some established techniques regarding uncertainty. However, the topic is – except for some specific cases (cf. Fig. 6.7) – mostly unexplored in the context of set visualizations. Some uncertainty visualization techniques may directly apply to set visualizations. In this seminar we brought together researchers from the areas of information visualization, visual analytics, graph drawing, geoinformation science, uncertainty research, and cognitive science. These interdisciplinary participants formed working groups to consider selected problems of considering and visualising uncertainty associated with sets so that the visualizations are informative and reliable, in the sense that humans can use them for visual analysis tasks and that the uncertain information is recognizable.

Seminar Format

The interdisciplinary topic of the seminar, as well as the different scientific backgrounds of the participants, asked for an introduction to the main topics as well as to selected perspectives through invited talks on the first day. The structure of two talks in the morning and two in the afternoon of the first day left enough room for first discussions. The day ended with participants' pitches of open problems and the participants indicating their interest in the pitched problems.

Invited talks of the first day:

- Daniel Archambault: Drawing Euler Diagrams with Closed Curves
- Wouter Meulemans: Algorithmic Perspectives on Uncertainty and Set Visualization
- Bei Wang Phillips: Visualizing Hypergraphs: With Connections to Uncertainty Visualization
- Martin Kryzwinski: Genomes: sets of sets of sets

The second day of the seminar was started with the formation of four groups interested in four different open problems. Each group worked on their specific open problem for the remainder of the seminar. Participants were invited to give mini-talks related to the seminar topic. Time was reserved for those contributed talks every morning. Additionally, the working groups reported on their progress on Wednesday and Friday.

Contributed mini-talks throughout the week (given are the names of the presenters, see Section 3 "Overview of Talks" in the full report for the full list of contributors):

- Annika Bonerath & Markus Wallinger: MosaicSets
- Sara Irina Fabrikant: How to visualize uncertainty
- Silvia Miksch: Visual Encodings of Temporal Uncertainty: A Comparative User Study
- Nathan van Beusekom: Simultaneous Matrix Orderings for Graph Collections
- Marc van Kreveld: On Full Diversity in Metric Spaces
- Alexander Wolff: StoryLines

Outcomes and Future Plans

The participants were highly satisfied with the quality of the seminar. Diverse interdisciplinary discussions took place and all groups worked well together. The final progress reports of the working groups indicate that the collaborations will be ongoing and some papers will be published (cf. Section "Working Groups" in the full report).

At the final day plenary meeting, plans for a follow-up seminar were discussed. A group of interested participants is currently discussing the focus and title of such a seminar.

References

1 Munzner, T. (2014). *Visualization Analysis & Design*. Boca Raton: CRC Press.

2 Padilla, L. and Kay, M. and Hullman, J. (2021). Uncertainty Visualization in *Wiley StatsRef: Statistics Reference Online*. Eds. N. Balakrishnan and T. Colton and B. Everitt and W. Piegorsch and F. Ruggeri and J.L. Teugels, Wiley. doi:10.1002/9781118445112.

3 Alsallakh, Bilal and Micallef, Luana and Aigner, Wolfgang and Hauser, Helwig and Miksch, Silvia and Rodgers, Peter (2016). The State-of-the-Art of Set Visualization. *Computer Graphics Forum*, 35(1):234–260. doi:10.1111/cgf.12722.

4 Steven Chaplick and Takayuki Itoh and Giuseppe Liotta and Kwan-Liu Ma and Ignaz Rutter (2020). *Trends and Perspectives in Graph Drawing and Network Visualization*, NII Shonan Meeting Report 171. https://shonan.nii.ac.jp/docs/342e2e86373d4783ae00063ddd67e8b35b9d75d5.pdf.

5 MacEachren, Alan M. and Robinson, Anthony and Hopper, Susan and Gardner, Steven and Murray, Robert and Gahegan, Mark and Hetzler, Elisabeth (2005). Visu-

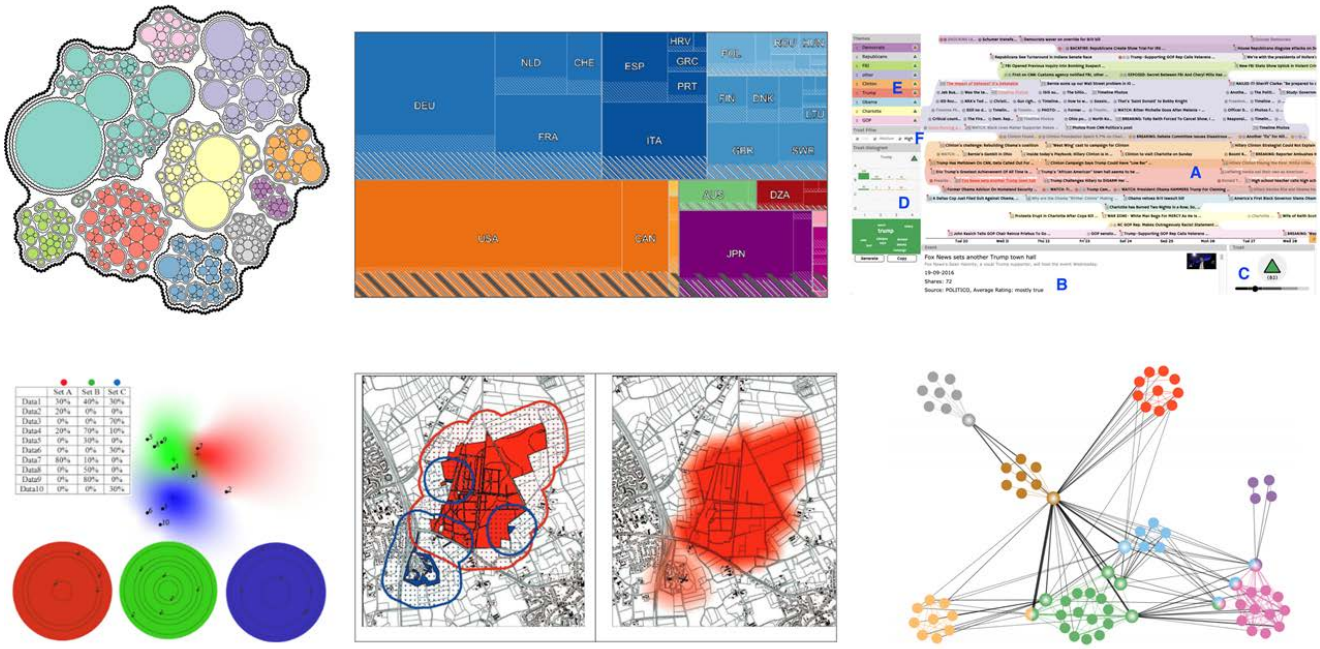


Fig. 6.7
Examples of uncertainty visualizations for sets from [6–11] (from top left to bottom right).

alizing Geospatial Information Uncertainty: What We Know and What We Need to Know. *Cartography and Geographic Information Science*, 32(3):139–160. doi:10.1559/1523040054738936.

6 Görtler, Jochen and Schulz, Christoph and Weiskopf, Daniel and Deussen, Oliver (2018). Bubble Treemaps for Uncertainty Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):719–728. doi:10.1109/TVCG.2017.2743959.

7 Sondag, Max and Meulemans, Wouter and Schulz, Christoph and Verbeek, Kevin and Weiskopf, Daniel and Speckmann, Bettina (2020). Uncertainty Treemaps. *2020 IEEE Pacific Visualization Symposium (PacificVis)*, 111–120. IEEE.

8 Xu, Kai and Salisu, Saminu and Nguyen, Phong H. and Walker, Rick and Wong, B. L. William and Wagstaff, Adrian and Phillips, Graham and Biggs, Mike and Potel, Mike (2020). TimeSets: Temporal Sensemaking in Intelligence Analysis. *IEEE Computer Graphics and Applications*, 40(3):83–93. doi:10.1109/MCG.2020.2981855.

9 Zhu, Lifeng and Xia, Weiwei and Liu, Jia and Song, Aiguo (2018). Visualizing fuzzy sets using opacity-varying freeform diagrams. *Information Visualization*, 17(2):146–160. doi:10.1177/1473871617698517.

10 Vullings, L. A. E. and Blok, C. A. and Wessels, C. G. A. M. and Bulens, J. D. (2013). Dealing with the uncertainty of having incomplete sources of geo-information in spatial planning. In *Applied spatial analysis and policy*, 6(1):25–45. Springer.

11 Vehlow, Corinna and Reinhardt, Thomas and Weiskopf, Daniel (2013). Visualizing Fuzzy Overlapping Communities in Networks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2486–2495. doi:10.1109/TVCG.2013.232.

6.72 Towards More Flexible and Automated Communication Networks

Organizers: Rui Aguiar, Artur Hecker, Stefan Schmid, and Henning Schulzrinne

Seminar No. 22471

Date: November 20–23, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.11.96

Participants: Gianni Antichi, Chen Avin, Roland Bless, Georg Carle, Klaus-Tycho Foerster, Fabien Geyer, Sergey Gorinsky, David Hay, Artur Hecker, Lily Hügerich, Karin Anna Hummel, Sándor Laki, Shir Landau Feibish, Huiran Liu, Gábor Rétvári, Dario Rossi, Iosif Salem, Gabriel Scalosub, Stefan Schmid, Henning Schulzrinne, Cigdem Sengul, Martina Zitterbart

Viewed from the perspective of users, communication networks just work fine and have not changed much recently. Under the hood, however, they are currently undergoing substantial changes, which are partly driven by new technologies, such as software-defined networking (SDN) and network function virtualization (NFV), and partly by new requirements, such as reducing operational costs and increasing reliability. SDN and NFV enable programming the behavior of these networks on-demand through software, so that functions and services can be flexibly deployed on short time scales at suitable locations in the network. The drawback of this enhanced flexibility is that, without proper tools, managing and operating such networks becomes more and more challenging. This complexity and the pressures to reduce costs call for largely autonomously operating – or self-driving – networks, i.e., networks with only limited manual intervention.

It is a challenge to provide robust and performant control planes and connectivity in such highly flexible and demanding networking environments, since the notions of control and data plane, by definition, are both related to the notion of the service to be provided. Hence, a network supposed to actively support services with challenging and varying requirements, will also be flexible, and this not just in terms of how it switches flows, but also in terms of the supported queuing models, protocol stacks, deployed and used network and service functions, and even in terms of its own topology. Such a flexible data plane will require an equally flexible control plane capable both of embracing new nodes with new capacities and capabilities and of re-allocating all tasks to new nodes in a shrinking network. The sharing of resources, both in capabilities and in capacities, needs to be efficiently supported not only between different services, but also between the respective control and data planes. Moreover, suitable distributed runtime scheduling algorithms are required in order to utilize and share network resources efficiently and to fulfill highly demanding requirements from certain network slices, e.g., ultra-reliable low latency communication in case of industrial networks. Also, network debugging and diagnostics need to cope with these new demands. It needs to be investigated, to what extent artificial intelligence and machine learning can be applied. This, however, is a rather new topic in the networking domain.

Consequently, the seminar brought together experts from "classical" networking, distributed systems and machine learning for networks.

6.73 Vertex Partitioning in Graphs: From Structure to Algorithms

Organizers: Maria Chudnovsky, Neeldhara Misra, Daniel Paulusma, and Oliver Schaudt

Seminar No. 22481

Date: November 27– December 2, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.11.109

Participants: Tara Abrishami, Akanksha Agrawal, Bogdan Alecu, Christoph Brause, Nick Brettell, Henning Bruhn-Fujimoto, Maria Chudnovsky, Konrad Dabrowski, Peter Gartland, Jan Goedgebeur, Petr A. Golovach, Bart Jansen, Tuukka Korhonen, Daniel Král', Madhumita Kundu, Paloma Lima, Daniel Lokshtanov, Barnaby Martin, Tomas Masarik, Jana Masarikova, Andrea Munaro, Jelle Oostveen, Sukanya Pandey, Daniel Paulusma, Marcin Pilipczuk, Michal Pilipczuk, Pawel Rzazewski, Saket Saurabh, Oliver Schaudt, Paul Seymour, Roohani Sharma, Siani Smith, Jan Arne Telle, Nicolas Trotignon, Erik Jan van Leeuwen, Kristina Vuskovic

Many important discrete optimization problems can be modelled as graph problems that ask if the set of vertices in a graph can be partitioned into a smallest number of sets, such that each set has the same property, or into some number of sets, such that each set has a specific property of their own. This leads to a rich framework of vertex partitioning problems, which include classical problems such as Graph Colouring, Graph Homomorphism, Vertex Cover, Feedback Vertex Set and Odd Cycle Transversal, and variants and generalizations of these problems.

Most vertex partitioning problems are computationally hard. The central research aim of our seminar was to increase our understanding of the computational complexity of these problems. The main approach followed at the seminar for achieving this was to restrict the input of some problem to some special graph class. The fundamental question then becomes whether such a restriction makes the problem tractable or whether the problem remains hard. In order to approach this question, we followed a systematic way by considering graph classes characterized by some finite family $\mathcal{H}$ of obstructions (as induced subgraph, subgraph, minor etc.).

In line with the seminar's research aim, the seminar brought together researchers from Discrete Mathematics, working in structural graph theory, and researchers from Theoretical Computer Science, working in algorithmic graph theory. In total, 36 participants from 12 different countries attended the seminar.

The scientific program of the seminar consisted of 23 sessions: 5 survey talks of fifty minutes, 13 contributed talks of at most thirty minutes and 5 open problem sessions. This left ample time for discussions and problem solving. Participants presented the progress that made during the workshop during several "progress report" sessions. One the the questions discussed at the workshop was a long-time open problem concerning the complexity of detecting whether a graph contains two long induced cycles with no edges between them. In the course of the workshop Khang Le informed us that he found a polynomial-time algorithm to solve this problem; Le's proof was presented by Paul Seymour.

Each of the five survey talks covered a particular structural or algorithmic key aspect of the seminar in order to enable collaborations of researchers with different backgrounds. On Monday, Paul Seymour presented a number of recent developments on the Erdös-Hajnal conjecture including several open problems. On the same day, Tara Abrishami described a variety of techniques for proving the boundedness or unboundedness of treewidth of hereditary graph classes. On Tuesday, Daniel Lokshtanov explained how algorithms for the Independent Set problem on P_k-free graphs developed over time, and also gave extensions of these results to other graph classes. On the same day, Daniel Král' surveyed basic results and open problems for two classical graph colouring parameters, the fractional and circular chromatic number of a graph, and one recent graph colouring parameter, the gyrochromatic number. On Wednesday, Henning Bruhn-Fujimoto gave a survey talk on Erdös-Pósa type questions, which relate to graph packing and graph covering dualities. In this talk, many open problems were given as well.

The five general open problem sessions took place on Monday, Tuesday, and Wednesday. Details of the presented problems can be found in the report, together with abstracts of all the talks.

We are grateful to Gerhard Woeginger for all his help with our seminar when it was originally planned to take place from January 31 to February 5, 2021. Our seminar was postponed to November 2022 due to the pandemic, and sadly, Gerhard passed away on April 1, 2022.

We also thank Akanksha Agrawal for her help with the Dagstuhl Report of our seminar.

6.74 Counting and Sampling: Algorithms and Complexity

Organizers: Holger Dell, Mark R. Jerrum, and Haiko Müller

Seminar No. 22482

Date: November 27– December 2, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.11.124

Participants: Konrad Anand, Nima Anari, Miriam Backens, Andreas Björklund, Marco Bressan, Andrei A. Bulatov, Sarah Cannon, Charlie Carlson, Amin Coja-Oghlan, Radu Curticapean, Ewan Davies, Holger Dell, Martin Dyer, Jacob Focke, Andreas Galanis, Andreas Göbel, Leslie Ann Goldberg, Heng Guo, Mark R. Jerrum, Petteri Kaski, John Lapinskas, Sarah Miracle, Haiko Müller, Noela Müller, Marcus Pappik, Viresh Patel, Guus Regts, Marc Roth, Philip Wellnitz, Stanislav Živný

Counting and sampling problems arise in areas such as statistics (benchmarking statistical tests, or sampling from a posterior distribution) and statistical physics (computing the partition function of a spin system). Computationally, these problems are very different in character from decision or optimisation problems, and their solution requires distinctive techniques. It is natural to treat counting and sampling together in the same Dagstuhl Seminar, as they are closely related computationally: subject to a reasonable side condition, an efficient algorithm for sampling certain combinatorial structures can be used as a black box to approximately count those structures, and vice versa.

Although much attention has been directed towards the complexity of counting and sampling problems, our understanding of them is not as well developed as it is of decision and optimisation problems. This seminar marks a timely return to the topic, as new ideas have recently been injected into the area, resulting in renewed activity and progress. It is particularly satisfying to observe that much of this progress has been in the positive direction, in the form of new efficient algorithms. This is in an area where negative results had become the norm.

The Covid pandemic inevitably left its mark on the meeting. Over five years elapsed between the previous Dagstuhl Seminar on a related topic and the current one. In the meantime, the introduction of a circle of ideas around high-dimensional expanders, spectral expansion and entropy decay has transformed the analysis of Markov chains for sampling, and brought many previously intractable questions within scope of our methods. An unwelcome impact of Covid was to reduce significantly the number of participants. Sadly, it was not possible to invite all the people we would have liked to see at the meeting.

With a view to providing a snapshot of current interests, here is a rough-and-ready breakdown of the presentations against a somewhat arbitrary set of headings.

- *Connections with statistical physics, phase transitions, etc.* Coja-Oghlan, Galanis and Patel,
- *Holant and constraint satisfaction problems.* Backens and Bulatov.
- *Markov chains.* Guo and Miracle.
- *Parameterised complexity of counting problems.* Bressan, Focke, Roth and Wellnitz.
- *Perfect samplers.* Anand and Cannon.
- *Point processes and other geometric connections.* Anari, Jerrum and Pappik.
- *Polynomials associated with graphs, matroids and matrices.* Björklund, Curticapean, Regts.
- *Other.* Göbel, Goldberg, Kaski, Lapinskas.

If nothing else, this rough classification exercise gives an impression of the wide span of current research. Aside from the progress on the analysis of Markov chains mentioned earlier, many other topics have seen advances in the past five years. Examples include: counting small patterns ('motifs') in large graphs (networks), sampling structures in regions of phase non-uniqueness, perfect sampling, and weighted counting problems where the weights are complex. It turns out that the latter study shines light on the case of real weights, through an examination of zeros of partition functions in the complex plane. The meeting gave participants a long-awaited chance to review developments over the past five years.

On the organisational front, an innovation (as far as this community is concerned) was the inclusion of a problem session on the first day. This went off quite smoothly, and small working groups formed fairly spontaneously to work on problems during the week. On the final day we heard from the groups a summary of their investigations over the week. Our hope is that sufficient momentum was achieved on some of these problems that groups will continue to work on them beyond the end of the meeting. Indeed, one of the working groups decided to apply to run a workshop on homomorphism counting at ICALP 2023 with this aim in mind. The proposal, by Radu Curticapean and Marc Roth, was accepted, and the workshop, entitled "ADjoint HOmomorphism Counting" (AD HOC) will take place in July 2023.

6.75 Cognitive Augmentation

Organizers: Kai Kunze, Pattie Maes, Florian 'Floyd' Mueller, and Katrin Wolf

Seminar No. 22491

Date: December 4–9, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.12.1

Participants: Michael Beigl, Michael D. Bonfert, Samantha W.T. Chan, Jiawen Han, Matthias Hoppe, Masahiko Inami, Shoya Ishimaru, Shunichi Kasahara, Marion Koelle, Thomas Kosch, Kai Kunze, Yuichi Kurita, Jie Li, Stephan Lukosch, Paul Lukowicz, Kouta Minamizawa, Qianqian Mu, Pat Pataranutaporn, Rakesh Patibanda, Roshan Lalintha Peiris, Enrico Rukzio, Albrecht Schmidt, Valentin Schwind, Paul Strohmeier, Steeven Villa, Tobias Wagner, Jamie A. Ward, Don Anusha Withanage, Katrin Wolf

Remote Participants: Cindy Hsin-Liu Kao, Zhuying Li, Pattie Maes, Florian 'Floyd' Mueller, Suranga Nanayakkara, Evangelos Niforatos, Nathalie Overdevest, Bektur Ryskeldiev, Aryan Saini, Stel Stelarc, Po-Yao (Cosmos) Wang

The real and digital worlds are increasingly more interconnected, leaving people to split their attention between tasks in the physical world in an increasing amount of ubiquitous systems and IoT services. We see an increase in accidents related to the usage of digital tools (such as interacting with a smartphone while driving). As governments and healthcare experts around the world call for changing lifestyles in response to the Covid-19 pandemic, the development, and usage of remote communication and touchless technologies are rapidly becoming an essential part of the "new normal". At the same time, the absence of touch and physical contact highlights their critical importance in human life, from school to hospital to care facilities. We need more intuitive, direct ways to interface with technology. Students and educators find it harder and harder to concentrate, and news outlets are already talking about the distraction economy. The seminar focused on people interacting with infor mation from the digital domain, in a minimally disruptive way, creating novel sensory experiences using and extending human perception and ultimately cognition.

The overall objective of the seminar was to foster research, explore, and model new means for increasing human intake of information in order to lay the foundation for augmented cognition, especially through somatosensation: the ability to sense the environment through our body.

Machine Learning has often been used to mimic or surpass some cognitive functions of the human mind (visual object/face recognition, playing chess, etc.). Such efforts appear to put humans and computers in a competitive relationship, as emphasized in AI vs. Human competitions. Once a fear of AIs "replacing" human workers is now taken much more seriously and discussed in the public sphere. This Dagstuhl Seminar suggests a different approach to the human-computer relationship by applying a cooperative and empowering framework. One important characteristic of the human mind is that it has significant fluctuations in productivity and capacity. Our mind has ebbs and flows, and is affected by various factors, some of which we do not even realize. These fluctuations manifest in patterns in human behavior and physiological signals (body temperature, eye movements, galvanic skin response, etc.). With this seminar, we wanted to discuss technologies that can give us more insights into the ebb and flow of the human mind as a basis for cognitive augmentation.

The participants developed several frameworks and taxonomies for understanding and evaluating different types of cognitive augmentation, based on their goals, methods, and impacts. The frameworks focus on enhancement, compensation, offloading, and replacement, and consider factors such as safety, efficacy, and social impact. There are several publication plans and several participants already agreed to organize conference workshops together (for example at AugmentedHumans 2023 and UbiComp/ISWC 2023).

Overall, the Dagstuhl Seminar on "Cognitive Augmentation" advanced our research fields by creating a shared understanding of the concept and its implications, promoting interdisciplinary collaboration and communication, and identifying promising directions for future research and development. The outcomes of the seminar are meant to inform the design, implementation, and evaluation of cognitive augmentation technologies, and contribute to augmenting human cognition by applying ethical principles.

6.76 Formal Methods and Distributed Computing: Stronger Together

Organizers: Hagit Attiya, Constantin Enea, Sergio Rajsbaum, and Ana Sokolova

Seminar No. 22492

Date: December 4–9, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.12.27

Participants: Hagit Attiya, Parosh Aziz Abdulla, Nathalie Bertrand, Raven Beutner, Annette Bieniusa, Ahmed Bouajjani, Manuel Bravo, Armando Castaneda, Gregory Chockler, Constantin Enea, Bernd Finkbeiner, Pierre Fraigniaud, Alexey Gotsman, Maurice Herlihy, Burcu Kulahcioglu Ozkan, Sandeep Kulkarni, Petr Kuznetsov, Ori Lahav, Giuliano Losa, Rupak Majumdar, Stephan Merz, Yoram Moses, Rotem Oshman, Azalea Raad, Sergio Rajsbaum, Ana Sokolova, Serdar Tasiran, Viktor Vafeiadis, Jennifer L. Welch, Thomas Wies, Philipp Woelfel

Distributed applications represent nowadays a significant part of our everyday life. To mention just a few examples, our personal data are stored on remote distributed servers, data management relies on remote applications reachable via smartphones or tablets, data-intensive computations are performed on computer clusters, etc. Since distributed applications are increasingly deployed at large scale, they have to be reliable and robust, satisfying stringent correctness criteria. This is the point where a strong interaction of formal methods and of distributed computing becomes a necessity.

The goal of this Dagstuhl Seminar was to achieve a synergy by bringing together researchers working on applying formal methods for concurrent programs and distributed systems, and researchers from distributed computing. Both communities have a deep understanding of distributed computation, but from two different perspectives. Historically, these communities have common roots, but since more than two decades they evolved independently. The resulting gap slows down progress in both fields, and limits the applicability of the results obtained in each field, as each one develops its own techniques separately. The seminar addressed several topics that bridge the two research fields, and that have high potential to stimulate the development of the other area:

Concurrent data structures and transactions: Modern multi-core architectures enable large performance boosts by executing a number of threads in parallel, which however, poses considerable challenges in maintaining correctness of shared data structures and thread synchronization. These challenges have been addressed using various paradigms like lock-free programming or transactional memory. However, turning these concepts into efficient programming support remains a big challenge, and formal methods may offer new ideas in this direction.

Formal approaches to large-scale replication: Current computing systems are increasingly large-scale distributed systems, for example, distributed databases, distributed ledgers (Blockchains) and key-value stores. At the heart of these systems are fundamental trade-offs between data consistency, availability, and the ability to tolerate failures. A formal approach to studying these issues will provide a common ground for the design, verification, analysis, implementation and use of these systems.

Distributed algorithms for verification: Reasoning about concurrent/distributed software is notoriously difficult due to the inherent non-determinism in its semantics. The different processes in a concurrent program can interleave in many different ways which leads to an enormous number of possible executions. Algorithmic methods are necessary to mitigate the difficulty of reasoning about this huge space of executions, and scalable distributed algorithms may be the answer for the future. These methods can manifest in various forms, e.g., automated testing, deductive verification, model checking, and have led to important results in many timely contexts. Performing verification in a distributed fashion is a particularly promising new direction of research.

The impact of all the areas above on a rigorous development of distributed applications was enhanced by fostering direct interactions between researchers from (automated) formal methods and from distributed computing.

6.77 Inter-Vehicular Communication – From Edge Support to Vulnerable Road Users II

Organizers: Ana Aguiar, Onur Altintas, Falko Dressler, and Gunnar Karlsson

Seminar No. 22512

Date: December 18–21, 2022 | Dagstuhl Seminar
Full report – DOI: 10.4230/DagRep.12.12.54

Participants: Ana Aguiar, Onur Altintas, Khalil Ben Fredj, Claudio Casetti, Carla-Fabiana Chiasserini, Klaus David, Falko Dressler, Jérôme Härri, Geert Heijenk, Frank Kargl, Gunnar Karlsson, Florian Klingler, Renato Lo Cigno, Marie-Christin Hannah Oczko, Jörg Ott, Michele Segata, Gürkan Solmaz, Christoph Sommer, Lukas Stratman, João P. Vilela, Lars Wolf

Looking back at the last decade, one can observe enormous progress in the domain of vehicular networking. In this growing community, many ongoing activities focus on the design of communication protocols to support safety applications, intelligent navigation, and many others. Using the terms "Vehicular Ad-hoc Networks" (VANETs), "Inter-Vehicle Communication" (IVC), "Car-2-X" (C2X), or "Vehicle-2-X" (V2X), many applications – as interesting as challenging – have been envisioned and (at least) partially realized. Very large projects have been initiated to validate the theoretic work in field tests and protocols are being standardized. With the increasing interest from industry, security and privacy have also become crucial aspects in the stage of protocol design in order to support a smooth and carefully planned roll-out. We are now entering an era that might change the game in road traffic management. Many car makers already supply their recent brands with cellular and WiFi modems, some also adding vehicular WLAN (DSRC, ITS-G5) and C-V2X technologies.

With this latest installment of the "Inter-Vehicular Communication" Dagstuhl Seminar series, we intend to shift the focus from basic networking principles to open challenges in edge computing support and, as a novel aspect, on how to integrate so called vulnerable road users (VRU) into the picture. Edge computing is currently becoming one of the core building blocks of cellular networks, including 5G, and it is necessary to study how to integrate ICT components of moving systems. The trade-offs of computation distribution, system aspects, and the impact on end-to-end latency are still unanswered. Also, vehicular networking and cooperative driving focus almost exclusively on cars but leave out communication and coordination with, for example, pedestrians and bicyclists. And, many of the existing communication solutions for this scenario were designed without having battery constraints in mind. In the meantime, some early research has been initiated on this topic and initial projects report very interesting results on safety features for VRUs. Building upon the great success of the previous Dagstuhl Seminars – as documented, e.g., with results published in widely visible magazine articles [1–4] – with this follow-up seminar, we wanted to again bring together experts from all these fields from both academia and industry.

Seminars in this series focused on general vehicular communication technologies, security and safety impact, cooperative driving concepts and its implications on communication protocol design, and many more. Building upon the online-only seminar in 2021, we now shifted the focus of this seminar from basic networking principles to open challenges in edge computing support and, as a novel aspect, on how to integrate so called vulnerable road users (VRU) into the picture. Edge computing is currently becoming one of the core building blocks of cellular networks, including 5G/6G, and it is necessary to study how to integrate ICT components of moving systems. The trade-offs of computation distribution, system aspects, and the impact on end-to-end latency are still unanswered. Also, vehicular networking and cooperative driving focuses almost exclusively on cars but leaves out communication and coordination with, for ex-ample, pedestrians and bicyclists. For example, many of the existing communication solutions for this scenario were designed without having battery constraints in mind.

The seminar focused intensively on discussions in several working groups. To kick-off these discussions, we invited four keynote talks:

- Vehicles and The Edge: Random thoughts and not so random Perspectives by Jörg Ott (TU Munich, DE)
- Who protects the Unprotected? ITS Services for Vulnerable Road Users by Claudio Casetti (Politecnico di Torino, IT)
- Enabling data spaces: Existing developments and challenges by Gürkan Solmaz (NEC, DE)
- Securing Cooperative Intersection Management by Subjective Trust Networks by Frank Kargl (Ulm University, DE)

We finally organized the following working groups on some of the most challenging issues related to inter-vehicular communication, edge computing, and vulnerable road users:

- Edge computing
- Vulnerable road users
- Vehicle to cloud to vehicle communication
- Sensing and analytics
- Trust

References

1 Falko Dressler, Frank Kargl, Jörg Ott, Ozan K. Tonguz and Lars Wischhof, "Research Chal-lenges in Inter-Vehicular Communication – Lessons of the 2010 Dagstuhl Seminar," IEEE Communications Magazine, vol. 49 (5), pp. 158–164, May 2011.

2 Falko Dressler, Hannes Hartenstein, Onur Altintas and Ozan K. Tonguz, "Inter-Vehicle Communication – Quo Vadis," IEEE Communications Magazine, vol. 52 (6), pp. 170–177, June 2014.

3 Onur Altintas, Suman Banerjee, Falko Dressler and Geert Heijenk, "Executive Summary – Inter-Vehicular Communication Towards Cooperative Driving," Proceedings of Dagstuhl Seminar 18202 on Inter-Vehicular Communication – Towards Cooperative Driving, Schloss Dagstuhl, Germany, May 2018, pp. 31–59.

4 Ana Aguiar, Onur Altintas, Falko Dressler and Gunnar Karlsson, "Executive Summary – Inter-Vehicular Communication – From Edge Support to Vulnerable Road Users," Proceedings of Dagstuhl Seminar 21262 on Inter-Vehicular Communication – From Edge Support to Vulnerable Road Users, vol. 11, Virtual Conference, June 2021, pp. 89–96.

7 Öffentlichkeitsarbeit
Public Relations and Outreach

7.1 Pressemitteilungen und Medienarbeit

7.1 *Press Releases and Media Work*

Die regelmäßige Erstellung und Herausgabe von Pressemitteilungen dient der verständlichen Verbreitung von aktuellen Informatikthemen. Die Vermittlung des Konzepts von Schloss Dagstuhl ist dabei ebenfalls ein Thema. Pressemitteilungen und Berichterstattungen in diversen Medien – soweit bekannt – sind über das Internetportal von Schloss Dagstuhl[48] abrufbar.

Schloss Dagstuhl hat sich zur allgemeinen Anlaufstelle für Journalisten etabliert, die über bestimmte Informatikthemen, aber auch über Schloss Dagstuhl berichten möchten. Durch Unterstützung des Saarländischen Rundfunks steht Schloss Dagstuhl ein professionelles Reporterset zur Verfügung, welches Rundfunkjournalisten erlaubt, vor Ort mit Seminarteilnehmern Interviews in digitaler, verlustfreier Audioqualität zu führen.

Schloss Dagstuhl verbreitet Neuigkeiten rund um sein Programm auch über soziale Netzwerkdienste wie Twitter und LinkedIn. Über Twitter-Nutzer @dagstuhl werden Programmankündigungen, die Publikation von neuen Tagungsbänden aber auch andere relevante Neuigkeiten an aktuell ca. 3 161 Abonnenten verbreitet. Zunehmend nutzen aber auch Seminarteilnehmer den Dienst, um ihre Eindrücke vom Seminar mitzuteilen. Darüber hinaus werden über den Twitter-Nutzer @dblp_org Informationen über die Bibliographiedatenbank dblp an ca. 1 890 Abonnenten verbreitet. Bei LinkedIn wird eine eigene Gruppe „Friends of Schloss Dagstuhl“ gepflegt (derzeit fast 620 Mitglieder), mit dem Ziel, die Vernetzung der Teilnehmer von Dagstuhl-Seminaren zu unterstützen.

Regular press releases showcase and disseminate information about current computer science topics in a comprehensible manner and clarify the concept behind Schloss Dagstuhl. Press releases and media reports that come to the center's attention are available on the Schloss Dagstuhl website[48].

Schloss Dagstuhl has become a port of call for journalists seeking to report on specific computer science topics and/or on Schloss Dagstuhl itself. Thanks to the support of the Saarländischer Rundfunk, Schloss Dagstuhl has access to professional reporting equipment that enables broadcast journalists to conduct interviews with seminar participants in digital lossless audio quality.

News on the program of Schloss Dagstuhl are also disseminated via social networks such as Twitter and LinkedIn. The Twitter handle @dagstuhl is used to disseminate program announcements, publication announcements, and other relevant news to about 3,161 followers, but is also increasingly used by Dagstuhl Seminar participants to share their impressions. Additionally, information about the dblp computer science bibliography is sent using the Twitter account @dblp_org, having about 1,890 followers. At LinkedIn, a “Friends of Schloss Dagstuhl” group is maintained (with almost 620 members), which supports the networking of participants in Dagstuhl Seminars.

7.2 Fortbildung

7.2 *Educational Training*

Lehrerfortbildung

Seit nunmehr 32 Jahren engagiert sich Schloss Dagstuhl im schulischen Bereich durch die Organisation einer jährlichen Lehrerfortbildung, die sich an Informatik- und Mathematiklehrer der gymnasialen Oberstufe im Saarland und in Rheinland-Pfalz richtet. Nach dem 2020 und 2021 die Pandemie die Durchführung verhinderte, gab es 2022 schließlich die 30. Lehrerfortbildung in Schloss Dagstuhl. Die Veranstaltung wird in Zusammenarbeit mit dem saarländischen Landesinstitut für Pädagogik und Medien (LPM) und dem Pädagogischen Landesinstitut Rheinland-Pfalz (PL) organisiert. Das LPM stützte die Fortbildung 2022 auch finanziell, indem es die Kosten der Referenten trägt.

Die Lehrerfortbildung dauert drei Tage; an jedem Tag

Teacher training

Since 32 years, Schloss Dagstuhl hosts an annual teacher training workshop specifically designed for teachers of upper secondary students working in the Saarland or the Rhineland Palatinate. After the pandemic stopped the plans for the training in 2020 and 2021, the 30th teacher training finally took place in Schloss Dagstuhl in 2022. The workshop is organized together with the Landesinstitut Pädagogik und Medien (LPM), Saarland, and the Pädagogisches Landesinstitut Rheinland-Pfalz (PL). The LPM supported the event finacially as well in 2022, by assuming the cost of the speakers.

The workshop lasts three days; each day two computer science topics are presented in a three hour presentation each. While this intensive training program mainly targets

[48] https://www.dagstuhl.de/institute/press

werden in jeweils 3-stündigen Vorträgen zwei Informatikthemen vorgestellt. Die intensive Fortbildung richtet sich zwar hauptsächlich an Lehrer aus dem Saarland und Rheinland-Pfalz, jedoch können seit 2011 bis zu fünf Lehrer aus anderen Bundesländern teilnehmen. Mehr Informationen zur Lehrerfortbildung 2022 gibt es auf der Webseite der Veranstaltung[49].

32 Jahre Informatik in Schloss Dagstuhl

Der 2020 wegen der Pandemie ausgefallene Festakt anlässlich 30 Jahre Informatik in Schloss Dagstuhl wurde als „2^5 Years of Computer Science at Schloss Dagstuhl“ am 24 Juni 2022 nachgeholt. Die rund 70 geladenen Gäste aus Politik und Wissenschaft lauschten nach den Grußworten der saarländischen Ministerpräsidentin Anke Rehlinger, der Ministerialdirektorin Ina Schieferdecker vom Bundesministerium für Bildung und Forschung und des Ministerialdirektors Daniel Stich vom Ministerium für Wissenschaft und Gesundheit Rheinland-Pfalz mehreren Festvorträgen: Die Vorsitzende des Wissenschaftsrates, Prof. Dr. Dorothea Wagner, legte nach einem Rückblick in die Geschichte des LZI auch die Herausforderungen der Gegenwart und der Zukunft dar. Prof. Dr. Meinard Müller von den International Audio Laboratories Erlangen gab einen Einblick in die Forschungen zur Musikinformatik, mit denen sich seine vergangenen Dagstuhl Seminare befasst hatten. Zum Abschluss stellte der Direktor der Initiative de Recherche et Innovation sur le Logiciel Libre, Prof. Dr. Roberto Di Cosmo, INRIAs Software Heritage Initiative zur Sammlung und Archivierung von Programmcode vor, deren Direktor er ebenfalls ist. An das Festkolloquium schloss sich zum Abrunden der Feierlichkeit ein Grillempfang im Garten und – leider wetterbedingt – im Speisesaals des Schlosses an. Das Program des Festakts findet sich auf der Webseite der Veranstaltung[50].

teachers from the Saarland and the Rhineland Palatinate, since 2011 up to five teachers of other federal states can participate. Details on the workshop in 2022 are available at the event webpage[49].

32 Years Informatics in Schloss Dagstuhl

The ceremony to mark 30 years of computer science at Schloss Dagstuhl, which was cancelled in 2020 due to the pandemic, was made up for with the Scientific Colloquium and Anniversary Ceremony "2^5 Years of Computer Science at Schloss Dagstuhl" on June 24, 2022. The approximately 70 invited guests from politics and science listened to welcoming addresses by Saarland Minister President Anke Rehlinger, Ministerial Director Ina Schieferdecker from the Federal Ministry of Education and Research and Ministerial Director Daniel Stich from the Ministry of Science and Health of Rhineland-Palatinate, followed by several keynote speeches: After a review of the history of the LZI, the Chair of the Science Council, Prof. Dr Dorothea Wagner, also outlined the challenges of the present and the future. Prof. Dr. Meinard Müller from the International Audio Laboratories Erlangen gave an insight into the research on music informatics that his past Dagstuhl Seminars had dealt with. Finally, the director of the Initiative de Recherche et Innovation sur le Logiciel Libre, Prof. Dr. Roberto Di Cosmo, presented INRIA's Software Heritage Initiative for the collection and archiving of programme code, of which he is also the director. To round off the festivities, the colloquium was followed by a barbecue reception in the garden and – unfortunately due to the weather – in the dining room of the castle. The program can be found on the event webpage[50].

[49] https://www.dagstuhl.de/22503
[50] https://www.dagstuhl.de/22255

8 Einrichtungen
Facilities

Das Zentrum verfügt über drei Standorte; der Hauptstandort ist Schloss Dagstuhl in Wadern. Die Geschäftsstelle mit Sachbearbeitungsteam und wissenschaftlichen Mitarbeitern, die für die Dagstuhl-Seminare und Perspektiven-Workshops verantwortlich sind, befinden sich auf dem Campus der Universität des Saarlandes in Saarbrücken, während der Bibliographiedienst durch Mitarbeiter vorwiegend in Räumlichkeiten der Universität Trier aber auch am Standort Saarbrücken betreut wird. Der Dagstuhl-Verlagsdienst befindet sich in Saarbrücken und Wadern.

The institute operates from three sites: the main site is Schloss Dagstuhl in Wadern. The administrative office and the scientific staff operating the Dagstuhl Seminars and Perspectives Workshops are located on the campus of Saarland University in Saarbrücken, while the staff operating the Bibliographic Services is mainly located in offices on the campus of the University of Trier but also in Saarbrücken. Dagstuhl Publishing is located in Saarbrücken and Wadern.

Hauptstandort in Wadern 8.1 *Main Site in Wadern*

Der Hauptstandort in Wadern umfasst das historische Schloss (gebaut um 1760) mit einem Anbau aus den 1970ern, einem 1993 fertiggestellten Erweiterungsbau, in dem sich Forschungsbibliothek, Hörsäle, Gästezimmer, Büros und Infrastruktur befinden, und ein 2012 fertiggestelltes Gästehaus mit Gästezimmern, einem Konferenzraum und Räumlichkeiten der Gebäudeverwaltung. Alle Einrichtungen in Wadern sind ganzjährig in Betrieb, abgesehen von je zwei Wochen im Sommer und Winter, die für größere Instandhaltungsarbeiten genutzt werden.

Die Kapazitäten von Dienstleistungen und Räumlichkeiten zur Veranstaltung von Seminaren sind genau aufeinander abgestimmt: Das Zentrum hat 71 Gästezimmer, davon sind 18 Doppelzimmer, sodass insgesamt 89 Teilnehmer übernachten können. Bei Normalbetrieb finden parallel zwei Seminare mit jeweils 30 und 45 Teilnehmern statt, wobei jedem Seminar ein Hörsaal für 35 bzw. 60 Personen zur Verfügung steht. Obwohl so eine Gesamtsumme von 75 Teilnehmern entsteht, ist es nur selten notwendig, Seminargäste in Doppelzimmern oder einem nahegelegenen Hotel unterzubringen. Die Obergrenze von 71 Zimmern wird regelmäßig (außerhalb von Pandemiezeiten) erreicht, weshalb es wohl kaum Möglichkeiten gibt, die Nutzung unserer Einrichtungen weiter auszubauen.

Mehr und mehr entwickelt sich aber das Konferenzzentrum in Wadern auch zu einem Treffpunkt aller Projekt- und wissenschaftlichen Mitarbeiter, die ansonsten getrennt über die verschiedenen Standorte bzw. im Homeoffice arbeiten. Hier können sie sich austauschen und gemeinsam vor Ort an den aktuellen Projekten und Aufgaben arbeiten.

The main site in Wadern comprises the historic manor house (built around 1760) with an extension from the 1970s; a facility completed in 1993, which is housing a research library, lecture halls, guest rooms, offices and infrastructure; and a guest house completed in 2012 with guest rooms, a conference room, and garages for facility management. All facilities at Wadern are operated all year round except for two weeks in summer and two weeks in winter when larger maintenance tasks are scheduled.

The capacities of services and facilities for hosting seminars at the main site are well coordinated: the site has 71 rooms, including 18 double rooms, for a total capacity of 89 participants staying overnight. During routine operation two seminars with nominally 30 and 45 participants are hosted in parallel, each using a lecture hall with 35 and 60 seats, respectively. Even though this sums up to 75 seminar participants, it is rarely necessary to book seminar guests into double rooms or a nearby hotel. The maximum capacity of 71 rooms is reached regularly (in non-pandemic times) and hence there is hardly a way to increase utilization of facilities further.

Incrementally, the conference center in Wadern is also developing into a meeting place for all project and scientific staff who are otherwise working in separate locations or in home office. Here, they can exchange ideas and work together on current projects and tasks.

Tagungsräume

Schloss Dagstuhl bietet drei Hörsäle für jeweils 25 bis 60 Personen. Alle Hörsäle sind auch für Hybridveranstaltungen ausgestattet. Darunter verstehen wir Veranstaltungen mit Teilnehmern die vor Ort teilnehmen und anderen, die per Videokonferenztechnik von ihrem perönlichen Rechner aus oder mittels der Technik eines der anderen Hörsäle teilnehmen. Die Technik dafür umfasst eine Audioanlage mit Lautsprechern und verschiedenen Mikrophonen. Neben einem Kopfbügelmikrophon und Handmikrophon sind mehrere Mikrophone an den Decken montiert, sodass alle Teilnehmer frei diskutieren können. Für die Videoübertragung stehen eine fernsteuerbare Kamera in Richtung des

Conference Facilities

Schloss Dagstuhl has three lecture halls with a seating capacity of 25 to 60 people each. All lecture halls are equipped for hybrid events. This entails that offered events can include a combination of on-site participants and those participating via video conference from their own personal computer or from one of the other lecture halls. The technology offered includes an audio system with speakers and various microphones. In addition to a headset microphone and a hand-held microphone, several microphones are mounted on the ceiling to foster discussions between all participants. For the video transmission, we offer a remote-controlled camera for the speaker and the content

Vortragenden und der Tafeln sowie eine fest ausgerichtete Kamera auf das Publikum zur Verfügung. Weiterhin sind jeweils ein zentraler Beamer für die Präsentation und ein weiterer Beamer bzw. ein 85 Zoll Display zur Ansicht der Remote-Teilnehmer vorhanden. Daneben steht in jedem Hörsaal auch ein MS-Windows-Laptop zur Verfügung.

Neben den Hörsälen gibt es im Zentrum sechs Seminarräume. Davon sind zwei mit modernen Beamern ausgestattet, während in einem ein großes Plasmadisplay montiert ist. Fünf Beamer auf Rollwagen stehen zusätzlich zur flexiblen Benutzung in allen Räumen zur Verfügung.

Alle Hörsäle und andere Tagungsräume sind mit Tafeln und/oder Whiteboards augestattet, selbst der Weinkeller bietet ein Whiteboard, damit Teilnehmer beim gemütlichen Zusammensein spontan ihre Ideen präsentieren können. Weitere mobile Flipcharts stehen flexibel für alle Veranstaltungen bereit. Die beiden größten Hörsäle sind jeweils mit mehreren Tafeln ausgestattet, während in den anderen Tagungsräumen jeweils große Whiteboards an den Wänden montiert sind.

Daneben gibt es über das ganze Zentrum verteilt weitere Räume, in denen Gäste sich in entspannter Atmosphäre treffen und diskutieren können. Insbesondere am Abend zieht es viele Gäste in den Weinkeller und die Cafeteria, zwei der gemütlichsten Räume im Haus und hervorragend geeignet für die Fortsetzung einer produktiven Diskussion in angenehmer Atmosphäre.

Dagstuhls Küche

Die Mahlzeiten sind ein wichtiger Bestandteil des wissenschaftlichen Programms von Schloss Dagstuhl. Die Sitzordnung wird absichtlich stets zufällig gemischt, um eingefahrene Gruppen aufzuteilen und Gäste zu ermuntern, während ihres Aufenthalts möglichst viele verschiedene Kollegen kennenzulernen. Große Tische im Speiseraum fördern die gemeinschaftliche Interaktion bei den Mahlzeiten.

Dagstuhls Philosophie des Kochens ist einfach: saisonal, gesund und schmackhaft. Unsere Gerichte werden jeden Tag von unseren Mitarbeitern der Küche frisch zubereitet. Der Schwerpunkt liegt dabei auf leichtem Essen während des Tages, um unsere Gäste nicht zu ermüden, und auf warmen Gerichten am Abend. Dies steht ein wenig im Widerspruch zur deutschen Tradition, kommt aber der Mehrheit der internationalen Gäste des Zentrums durchaus entgegen.

Sowohl die Zutaten als auch die Gerichte wechseln saisonal. An warmen Sommerabenden wird auf Anfrage auf der Terrasse vor dem Speisesaal gegrillt, unter anderem saarländische Schwenker, eine lokale Variante des Grillsteaks, die unter dauerndem Schwenken des Grillrostes zubereitet wird. In den kalten Monaten steht einmal wöchentlich ein schmackhafter Eintopf auf dem Speiseplan. Über das Jahr hinweg wird eine ausgewogene Mischung an regionalen und internationalen Spezialitäten aus neuen sowie bewährten und beliebten Rezepten angeboten. Im Allgemeinen sind die angebotenen Gerichte im Sommer etwas leichter und im Winter ein wenig schwerer. Die Küche arbeitet nach dem HACCP-Konzept (Hazard

of the boards, as well as a fixed camera capturing the on-site audience. Furthermore, there is a projector for the speakers presentation and an additional projector or an 85 inch display showing the remote participants. A laptop with MS-Windows is available in every lecture hall.

In addition to the lecture halls, the center has six meeting rooms. Two are equipped with up-to-date projectors and one has a large plasma display on the wall. Five mobile projectors are available for use in all of the rooms.

All lecture halls and meeting rooms are equipped with blackboards and/or whiteboards. Even the wine cellar offers a whiteboard, so participants can spontaneously pitch their ideas during a relaxing get-together. Additional mobile flipcharts are available for all events. Whereas the two main lecture halls are equipped with several blackboards, whiteboards are provided in the other rooms.

The center also offers a variety of other spaces where guests can sit and work together in a relaxed atmosphere. Particularly in the evening, guests gravitate towards the wine cellar and upstairs café, two of the coziest places in the house and great places for continuing a productive discussion in a comfortable atmosphere.

Dagstuhl's Kitchen

The dining experience at Dagstuhl is an important part of the center's scientific program. Seating arrangements are mixed deliberately in order to break up cliques and encourage guests to talk to as many different people as possible during the course of their stay. Large tables in the dining hall promote collaborative interaction during meals.

The philosophy behind Dagstuhl's cooking is simple: seasonal, healthy, and tasty meals. Everything is freshly prepared each day by the kitchen's staff. The focus is on lighter fare during the day in order to aid scientists' concentration, and on a warm meal in the evening, breaking with the German tradition of a cold evening meal while matching the internationality of the center's guests.

Both ingredients and dishes vary with the seasons. On warm summer evenings, guests are invited on demand to partake of grilled *Schwenker* (the local variant of barbecued steak) on the outdoor patio adjacent to the dining hall. During the cold winter months, warm soups appear on the menu weekly. In general, the kitchen tries to keep meals lighter in the summertime and heavier in the winter, offering a blend of regional and international dishes year-round that include some new recipes and many tried-and-true Dagstuhl favorites. The kitchen works in accordance with the HACCP Concept (Hazard Analysis and Critical Points Concept) and adheres to the mandatory labeling of allergens, which is required of all food processing establishments. Food additives and conservatives for which labeling is non-mandatory are also carefully monitored.

All guests with special dietary requirements due to ethical or health reasons can announce their needs prior to

Analysis and Critical Points Concept) und hält sich an die Kennzeichnungspflicht von Allergenen, zu der alle lebensmittelverarbeitenden Betriebe verpflichtet sind. Des Weiteren achten wir auf deklarationsfreie Zusatz- und Konservierungsstoffe.

Alle Gäste, die aus medizinischen oder ethischen Gründen Einschränkungen bei der Speiseauswahl haben, können sich vor dem Seminar bei Schloss Dagstuhl melden. Unsere Küchenmitarbeiter erarbeiten gerne individuelle Lösungen für jeden Gast, soweit es irgend möglich ist. Gäste, die koscheres Essen benötigen, haben die Möglichkeit, mitgebrachte abgepackte Speisen selbst zu erhitzen.

Um unseren Gästen trotz eines begrenzten Budgets eine ausgewogene Qualität anbieten zu können, bietet unsere Küche ein Frühstücksbüffet, dienstags bis donnerstags abhängig von den personellen Kapazitäten ein Mittagsbuffet sowie ein Menü am Abend an. Montags und freitags wird aus logistischen Gründen auch am Mittag ein Menü serviert. Unser Restaurant mit den großen Fenstern zum Garten des Hauptgebäudes bietet ca. 80 Personen Platz. Hier herrscht eine entspannte und fast familiäre Atmosphäre, was nicht zuletzt auf unsere freundlichen und engagierten Mitarbeiter zurückzuführen ist.

Kleine und große Pausen unterbrechen auf angenehme Weise die tägliche Routine und anstrengenden Diskussionen. In der kleinen Kaffeepause am Vormittag stehen vor den Vortragsräumen heiße Getränke auf einem Kaffeewagen bereit. In der großen Kaffeepause am Nachmittag wird den Gästen im Speiseraum neben heißen Getränken auch frisch gebackener Kuchen angeboten. Darüber hinaus gibt es im Gästehaus, der „alten“ Cafeteria und dem Weinkeller jeweils einen Kaffeevollautomaten zur Zubereitung von Kaffee, Kakao und Tee. Im Kiosk vor der Cafeteria können Gäste Snacks erwerben. Abends gibt es in der Cafeteria und im sogenannten Weinkeller einen Gruß aus der Küche, bestehend aus Brot und einer Käseauswahl.

Kinderbetreuung

Schloss Dagstuhl bietet Teilnehmern, die mit Kindern anreisen müssen, ein qualifiziertes Betreuungsprogramm für Kinder an. Dieser von Schloss Dagstuhl subventionierte Service kann gegen ein geringes Entgelt im Voraus gebucht werden. Alternativ ist es Eltern auch möglich, eine Begleitperson zur Betreuung des Kindes oder der Kinder mitzubringen. Schloss Dagstuhl kommt für die Unterkunft und Verpflegung der Kinder auf. Wenn statt Inanspruchnahme der Kinderbetreuung von Schloss Dagstuhl eine Betreuungsperson mitreist, hat diese ebenfalls freien Aufenthalt. Schloss Dagstuhl hält einen Aufenthaltsraum für Kinder vor. Dieser ist unter anderem mit einer Sitzgruppe, einem Spielhaus, einer Spielhöhle und Kinderzimmerraumteilern, die etwa mit einer Kugelbahn oder einem Spiegel ausgestattet sind, versehen. Er erlaubt einen angenehm Aufenthalt der Kinder auch im Innenraum.

Dagstuhls Angebot der Kinderbetreuung für Eltern wird gut genutzt. Obwohl im ersten Quartal keine Kinderbetreuung stattfinden durfte, wurde das Angebot im Jahr 2022 rege genutzt. Insgesamt 21 Teilnehmern von 16 Veranstaltungen mit 25 Kindern konnte eine Seminarteilnahme

the events. Our kitchen staff will then work out individual solutions if at all possible. Guests who need kosher meals can heat up ready-to-eat meals for themselves.

To accomplish all of this within a reasonable budget, the center offers a buffet-style breakfast and a set evening meal served by the kitchen's friendly and dedicated staff. From Tuesday to Thursday the kitchen offers a buffet-style lunch depending on the staff capacities. Due to logistical reasons, a set meal is served at lunch on Mondays and Fridays. The large dining-hall, seating up to 80 people, opens onto the castle garden and patio, and offers a relaxed, familiar atmosphere.

Small and late-morning breaks break up the daily routine. During the small coffee break in the morning, hot drinks are served outside the lecture halls. During the longer coffee break in the afternoon, hot drinks together with freshly baked cake are served in the dining hall. In addition, there are self-service bean-to-cup coffee machines in the guest house, at the "old" café, and in the wine cellar. Guests can buy small snacks at the kiosk in front of the café. Bread and cheese is served in the café and the wine cellar every night.

Childcare

Schloss Dagstuhl offers participants who need to travel with children childcare options. A certified nanny service, which is subsidized by Schloss Dagstuhl, can be booked in advance for a small fee. Alternatively, parents can bring an accompanying person to look after their child or children and when the nanny service is not used, this person stays free of charge. In both cases, Schloss Dagtuhl covers the cost of accommodation and meals of the children. Schloss Dagstuhl has a recreation room for children. It includes a seating area, a playhouse, a play cave, and children's room dividers equipped with a marble run, a mirror, and so on. This room gives the children a pleasant place to play indoors as well.

Dagstuhl's offered childcare services for parents are well used. Although there was no childcare in the first quarter, the offer was actively used in 2022. A total of 21 participants from 16 events with 25 children were able to attend a seminar thanks to the childcare. Of these, 7 children were cared for by Schloss Dagstuhl's nanny.

durch die Kinderbetreuung ermöglicht werden. 7 Kinder davon wurden durch Schloss Dagstuhls Kinderbetreuerin betreut.

Freizeit und Ambiente

Die Freizeitanlagen auf Schloss Dagstuhl wurden so gestaltet, dass sie auf unterschiedliche Art und Weise sowohl tagsüber als auch abends die Kommunikation zwischen den Seminarteilnehmern fördern. Die Mischung aus Arbeit und Freizeit in entspannter, familiärer Atmosphäre ist ein wichtiger Bestandteil des Dagstuhl-Konzepts. Gäste leben und arbeiten zusammen in einem Komplex aus drei Gebäuden, im Zentrum das historische Schloss, wo sie rund um die Uhr freien Zugang zu den zahlreichen Freizeiträumen und -anlagen haben. Musikalische Gäste können ihre Fertigkeiten im barocken Musiksaal zu Gehör bringen, wo ein Flügel und diverse andere Instrumente wie z. B. zwei Konzertgitarren zur Verfügung stehen. Unser Zentrum verfügt außerdem über eine Sauna, einen Billardtisch, Tischfußball, Mountainbikes, eine Dartscheibe, einen Freizeitraum mit Fitnessgeräten und Tischtennis sowie einen Außenbereich mit Volleyballnetz.

Leisure Facilities

Leisure facilities at Schloss Dagstuhl are designed to encourage and support communication among seminar participants in different settings throughout the day and evening. This work/life continuum within a relaxed, informal setting is an important part of the Dagstuhl concept. Guests live and work together in a complex of three buildings, the historical manor house ("Schloss") in the middle, and enjoy full access to the center's many unique rooms and facilities around the clock. Musically talented guests are welcome to exercise their skills in the baroque music room on the upper floor of the historical main building, which features a grand piano and various other instruments, e.g., two concert guitars. Schloss Dagstuhl also has a full sauna, a pool table, table football facilities, mountain bikes, a dartboard, and a recreation room with gym equipment and table tennis as well as outdoor sports grounds featuring a volleyball net.

8.2 Geschäftsstelle in Saarbrücken

8.2 *Dagstuhl Office in Saarbrücken*

Die Geschäftsstelle in Saarbrücken befindet sich auf dem Campus der Universität des Saarlandes im Gebäude E11. Die Räumlichkeiten werden vom Sachbearbeitungsteam und von dem Teil des wissenschaftlichen Stabs genutzt, der für die Seminarbetreuung zuständig ist. Zusammen mit dem Wissenschaftlichem Direktor Raimund Seidel, der einen Lehrstuhl an der Universität des Saarlandes in Saarbrücken hat, sind so alle, die bei der Gestaltung des Wissenschaftlichen Programms eine wesentliche Rolle spielen, an einem Standort.

In den Räumen der Geschäftsstelle arbeiten aber auch Mitarbeiter des Verlagsservices Dagstuhl Publishing und der Bibliographiedatenbank dblp, die teils auch Stellen in mehr als einem der Bereiche von Dagstuhl haben.

The Dagstuhl Office in Saarbrücken is located on the campus of Saarland University in building E11. The site houses some administrative staff and the part of the scientific staff responsible for seminar support. Together with the Scientific Director, Raimund Seidel, who holds a chair at Saarland University, all those who play an essential roll in the design of the scientific program are in one location.

However, employees of Dagstuhl's publishing service and bibliography database also work at this location, some of whom also have positions in more than one of Dagstuhl's three areas.

8.3 Dagstuhl an der Universität Trier

8.3 *Dagstuhl at the University of Trier*

Die meisten für die Bibliographiedatenbank dblp zuständigen Mitarbeiter haben ihren Standort an der Universität Trier. Die Ende 2010 zunächst auf Basis zweier Projekte gestartete Zusammenarbeit zwischen Schloss Dagstuhl und der Universität Trier wurde im November 2018 in eine offizielle und permanente Außenstelle von Schloss Dagstuhl auf dem Campus der Universität Trier überführt. Dabei profitiert das dblp-Team von der engen Zusammenarbeit mit der Abteilung Informatikwissenschaften und als externer Partner im Digital Research and Bibliographic Meta Data Lab des Center for Informatics Research and Technology (CIRT).

The scientific and editorial staff working on the *dblp computer science bibliography* is mainly located at the Dagstuhl offices at the University of Trier. Initially based on a project-based cooperation between Schloss Dagstuhl and the University of Trier which was first established in 2010, in November 2018, an official and permanent Schloss Dagstuhl branch office has been established on the campus of the University of Trier. In Trier, the dblp team benefits from the close cooperation with the University's department of computer science, and as an external partner in the Center for Informatics Research and Technology (CIRT) lab for Digital Research and Bibliographic Meta

Data.

9 Zentrale Dienste
Central Services

Schloss Dagstuhl verfügt über zwei zentrale Dienste: die IT-Abteilung und eine Forschungsbibliothek. Beide Einrichtungen befinden sich am Hauptstandort in Wadern.

Bibliothek

Zur wissenschaftlichen Literatur- und Informationsversorgung der Seminarteilnehmer unterhält Schloss Dagstuhl eine hervorragende Forschungsbibliothek für Informatik.

Die Bibliothek ist für Wissenschaftler vor Ort rund um die Uhr und für externe Wissenschaftler nach Absprache zugänglich. Zur digitalen Informationsinfrastruktur gehören ein Online-Bibliothekskatalog, ein modernes Discovery-System zur Artikelrecherche sowie zahlreiche Angebote für den Online-Zugriff auf wissenschaftliche Publikationen.

Für jedes Seminar wird eine individuelle Buchausstellung zusammengestellt, bestehend aus Büchern, die von Seminarteilnehmern verfasst oder herausgegeben wurden. Die anwesenden Autoren werden gleichzeitig gebeten, ihre Bücher zu signieren. Zur Optimierung der Autorenidentifikation werden die ORCIDs der Personennamen im Bibliothekskatalog erfasst.

Außerdem wird der Name eines jeden Seminarteilnehmers in der Online-Teilnehmerliste mit seinen oder ihren in der dblp-Literaturdatenbank erfassten Veröffentlichungen verlinkt. Diese Maßnahmen ermöglichen den Seminarteilnehmern einfachen und schnellen Zugriff auf seminarrelevante Literatur.

Die Bibliothek verfügt über einen umfangreichen Buchbestand, der Zugriff auf aktuelle Forschungspublikationen wie Konferenzbände und wissenschaftliche Zeitschriften erfolgt ausschließlich digital.

- Der Buchbestand orientiert sich am wissenschaftlichen Seminarprogramm. Bei Neuanschaffungen liegt der Fokus auf Büchern, die einen Bezug zu Dagstuhl-Seminaren oder Perspektiven-Workshops haben oder von Seminarorganisatoren oder -teilnehmern verfasst wurden. Außerdem erhält die Bibliothek zahlreiche Bücher als Spenden von Verlagen und Autoren. Aktuell verfügt die Bibliothek über etwa 36 000 Informatikbücher. Die Metadaten werden standardisiert erfasst und mit Hyperlinks angereichert, die durch persistente Adressierung (DOIs) verlässlich verlinkt sind.
- Beiträge in Konferenzbänden verkörpern den wichtigsten Teil der Literatur in der Informatik. Die Bibliothek hat die kompletten ACM- und IEEE-Proceedings elektronisch abonniert. Ältere Bände stehen teilweise auch in Druckform zur Verfügung. Die Verlagsgruppe Springer Nature spendet der Bibliothek alle Bände der Reihe Lecture Notes in Computer Science (LNCS) in Druckform. Die Bibliothek verfügt somit über Druckexemplare aller veröffentlichten Bände ab dem Band 1.
- Wissenschaftliche Fachzeitschriften sind eine wesentliche Voraussetzung für exzellente Forschung. Häufig werden in Zeitschriften erweiterte Fassungen von

Schloss Dagstuhl has two central services: the IT service and a research library, which are both located at the main site in Wadern.

9.1 *Research Library*

Schloss Dagstuhl maintains an excellent research library for computer science to provide seminar participants with scientific literature and information.

The library is accessible to on-site researchers around the clock and to external researchers by appointment. The digital information infrastructure includes an online library catalog, a modern discovery system for article research, as well as numerous options for online access to scientific publications.

For each seminar, an individual book exhibition is compiled, consisting of books written or edited by seminar participants. The authors who are present at the seminar are asked to sign their own books. In order to optimize the author identification, the ORCIDs of the authors' names are recorded in the library catalog.

In addition, the name of each seminar participant will be linked in the online list of participants with their publications recorded in the dblp literature database. These measures provide seminar participants with easy and quick access to the literature relevant to the seminar.

The library maintains an extensive collection of books. Access to current research publications such as conference proceedings and scientific journals is exclusively digital.

- The book collection is oriented towards the scientific seminar program. New acquisitions focus on books which are related to Dagstuhl Seminars and Perspectives Workshops or which were written by seminar organizers or participants. In addition, the library receives numerous books as donations from publishers and authors. Currently, the library has about 36,000 books on computer science. The metadata are recorded in a standardized way and enriched with hyperlinks, which are reliably linked by permanent addressing (DOIs).
- Contributions in conference proceedings represent the most important part of the literature in computer science. The library has subscribed to the complete ACM and IEEE proceedings electronically. Earlier volumes are also partly available in printed form. The Springer-Nature publishing group donates all printed volumes of the series Lecture Notes in Computer Science (LNCS) to the library. The library thus has print copies of all published volumes from volume 1 onwards.
- Scientific journals are essential for excellent research. Journals often publish extended versions of results that were previously published in conference proceedings. The library provides access to several thousand digital scientific journals. Most of them are included in journal packages licensed in cooperation with nationwide consortia, such as DFG-funded national and alliance licenses, as well as consortium licenses funded by the

Ergebnissen veröffentlicht, die zuvor in Konferenzbänden publiziert wurden. Die Bibliothek bietet Zugriff auf mehrere Tausend digitale Fachzeitschriften. Die meisten sind in Zeitschriftenpaketen enthalten, die in Kooperation mit deutschlandweiten Konsortien lizenziert sind, beispielsweise DFG-geförderte National- und Allianzlizenzen, Projekt DEAL, sowie von der Leibniz-Gemeinschaft geförderte Konsortiallizenzen.

- Die Bibliothek ermöglicht den benutzerfreundlichen Online-Zugriff auf über 7 000 deutschlandweite und internationale Zeitungen und Magazine aus über 120 Ländern.

Leibniz Association, and Projekt DEAL.

- The library enables user-friendly online access to over 7,000 Germany-wide and international newspapers and magazines from over 120 countries.

Zusammenarbeit

Schloss Dagstuhls Forschungsbibliothek ist mit zahlreichen überregionalen Bibliotheksdatenbanken vernetzt. Der komplette Zeitschriftenbestand ist in der Zeitschriftendatenbank (ZDB) nachgewiesen. Zusätzlich ist der Bestand an elektronischen Zeitschriften in der kooperativen bundesweiten Elektronischen Zeitschriftenbibliothek (EZB) erfasst. Darüber hinaus wird der komplette Monographienbestand im K10plus, der gemeinsamen Katalogisierungsdatenbank von GBV und SWB mit über 180 Millionen Nachweisen, nachgewiesen.

Diese Datenbanken bilden die Grundlage für den deutschlandweiten und internationalen Leihverkehr der Bibliotheken. Somit steht der Zeitschriftenbestand auch standortübergreifend und überregional für Fernleihzwecke zur Verfügung.

Außerdem besteht eine enge Zusammenarbeit zwischen Schloss Dagstuhl und der Saarländischen Universitäts- und Landesbibliothek (SULB), der Campusbibliothek für Informatik und Angewandte Mathematik an der Universität des Saarlandes sowie der Bibliothek des Leibniz-Instituts für Neue Materialien (INM), die sich alle in Saarbrücken befinden.

Schloss Dagstuhls Fachbibliothek ist institutionelles Mitglied des Deutschen Bibliotheksverbandes (DBV). Die Bibliothekarin Petra Meyer ist persönliches Mitglied im Berufsverband Information Bibliothek e.V. (BIB).

Collaboration

Schloss Dagstuhl's research library is connected to numerous national library databases. The complete journal inventory is recorded in the Zeitschriftendatenbank (ZDB). In addition, the inventory of electronic journals is recorded in the cooperative nationwide Electronic Journals Library (Elektronische Zeitschriftenbibliothek, EZB). Furthermore, the complete Monograph collection is recorded in K10plus, the joint cataloging database of GBV and SWB with over 180 million records.

These databases form the foundation for the libraries' nationwide and international lending system. Thus the journal collections are also available for inter-library loan purposes across locations and regions.

There is also a close cooperation between Schloss Dagstuhl and the Saarland University and State Library (SULB), the Campus Library for Computer Science and Applied Mathematics at Saarland University, and the library of the Leibniz Institute for New Materials (INM), all of which are located in Saarbrücken.

Schloss Dagstuhl's specialized library is an institutional member of the German Library Association (Deutscher Bibliotheksverband, DBV). The librarian, Petra Meyer, is a personal member of the Professional Association Information and Libraries (Berufsverband Information Bibliothek e.V., BIB).

Spenden an die Bibliothek

Die Bibliothek von Schloss Dagstuhl profitiert von zahlreichen Spenden. So erhielt die Informatik-Fachbibliothek im Jahr 2022 Buchspenden vom Springer Verlag und auch von Seminarteilnehmern. Autorenexemplare werden ebenso dankbar entgegengenommen. Insgesamt erhielt das Zentrum im Berichtszeitraum 569 Bände, darunter 553 Monographien des Springer Verlags.

Library Donations

The Dagstuhl Informatics Research Library benefits from numerous book donations. In 2022, the Informatics Research Library received book donations from Springer Verlag and also from seminar participants. The center is also grateful for donations of author's copies. The center received a total of 569 volumes during the reporting period, including 553 monographs from Springer Verlag.

IT-Service 9.2 *IT Service*

Die IT-Abteilung bietet umfassenden Support für alle internen Vorgänge an den drei Standorten. Darüber hinaus betreut sie die IT-Infrastruktur und -Dienste und bietet Unterstützung für alle Gäste bei Dagstuhl-Veranstaltungen.

Der IT-Service umfasst u.a.:

The IT service provides comprehensive support for all internal operations at all three sites. Moreover, it provides IT infrastructure, services, and support for all guests of Dagstuhl events.

This service includes – among others – the following:

- Internetzugang über Ethernet und WLAN in allen Räumen. Für den WLAN-Zugang bietet Schloss Dagstuhl persönliche Accounts an und ist auch an der *eduroam*-Initiative beteiligt[51] (eine praktische Alternative für Gäste, die bereits einen *eduroam*-Account haben). Innerhalb sämtlicher Einrichtungen stellt Schloss Dagstuhl ein weitläufiges Netzwerk von Zugangspunkten zum Drahtlosnetzwerk zur Verfügung, das aktiv überwacht und regelmäßig erweitert wird. Die Verbindung zum (externen) Internet wird durch zwei redundante 375 Mbit/s-Leitungen sichergestellt, betrieben durch den DFN e.V. (Deutsches Forschungsnetz).
- Fahrbare ebenso wie fest montierte Präsentationsmöglichkeiten in den Tagungsräumen. In den größeren Tagungsräumen können Vortragende den vorhandenen oder den eigenen Laptop verwenden.
- Zwei Bildschirmarbeitsplätze in der Bibliothek, hauptsächlich zu der Recherche von Literatur.
- Ein Terminal vor der Rezeption, das den Check-in und das Schreiben von Ersatzzugangskarten zu den Gästezimmern erlaubt, auch wenn die Rezeption gerade unbesetzt ist.
- Ein Terminal vor der Cafeteria, das den schellen Ausdruck von Tickets erlaubt. Dieses Terminal kann im Bedarfsfall auch für das Schreiben von Zimmerzugangskarten benutzt werden.
- Zugang zu Netzwerkfarbdruckern, einem Scanner und einem Kopierer.
- Technischen Support für Seminarteilnehmer und Mitarbeiter von Schloss Dagstuhl.

Der IT-Service verwaltet (virtuelle) Server für alle Abteilungen, z.B.

- einen Webserver, auf dem sich Schloss Dagstuhls Internetpräsenz befindet (https://www.dagstuhl.de), die Informationen für Teilnehmer, zum Seminarprogramm usw. enthält,
- einen Server, auf dem sich Schloss Dagstuhls Publikationsplattform DROPS (https://drops.dagstuhl.de) befindet und
- den dblp-Server (https://dblp.dagstuhl.de oder, alternativ, https://dblp.org).

Darüber hinaus stellt der IT-Service Tools für das gemeinschaftliche Arbeitsumfelds zur Verfügung und hält sie in Stand, z.B. *Sihot* (eine Software zur Organisation von Gastdaten), MySQL-Datenbanken, ownCloud (ein Cloud-basiertes Speichersystem) und weitere.

- Internet access via Ethernet and Wi-Fi throughout all rooms. For Wi-Fi access, Schloss Dagstuhl offers personal accounts and also takes part in the *eduroam* service[51] (which is a comfortable option for guests with existing *eduroam* accounts). Within its facilities, Schloss Dagstuhl provides a vast network of professional-grade wireless network access points that is actively monitored and extended regularly. External internet access for Schloss Dagstuhl is provided through two redundant 375 Mbit/s connections that are managed by DFN e.V. (National Science Network).
- Mobile and stationary presentation facilities in meeting rooms. In large meeting rooms, presenters can use either a provided laptop or their own.
- Two workstations in the library, mainly for literature research purposes.
- One terminal in front of the reception, allowing for self check-in and the writing of replacement room cards, even if there is no one at the reception desk.
- A terminal in front of the coffee lounge that can be used to print tickets. This terminal can also be used to write room cards.
- Access to network color printers, a scanner, and a photocopier.
- Technical support for both seminar participants and Dagstuhl staff.

The IT service manages (virtualized) servers for Schloss Dagstuhl's divisions, such as

- a web-server hosting Schloss Dagstuhl's web page at https://www.dagstuhl.de, providing information for participants, information about the seminar program, etc.,
- a server hosting DROPS at https://drops.dagstuhl.de, Schloss Dagstuhl's publishing platform, and
- the dblp server at https://dblp.dagstuhl.de or, alternatively, at https://dblp.org.

Furthermore, for internal work procedures, the IT service provides and maintains tools for a collaborative work environment, such as *Sihot* (a software for organizing guest data), MySQL data bases, ownCloud (a cloud-based storage system), and several others.

[51] *eduroam* (education roaming) is a world-wide roaming access service developed for the international research and education community, see https://www.eduroam.org.
eduroam *(education roaming) ist ein weltweiter roaming Zugangsdienst der für die internationale Forschungs- und Bildungsgemeinde entwickelt wurde, siehe https://www.eduroam.org.*

10 Kunst
Art

Dagstuhl als Galerie 10.1 *Dagstuhl as an Art Gallery*

Im sogenannten Kreuzgang des Neubaus werden regelmäßig Kunstausstellungen organisiert. Das großzügige Raumangebot der Wände des Flurs sowie die hervorragende Ausleuchtung mit starken Kontrasten zwischen Tag und Nacht bieten den Künstlern sehr gute Möglichkeiten, ihre Werke darzustellen. Die Kunstwerke an den Wänden des schmalen Gangs durchbrechen die Nüchternheit des Neubaus in anregender und angenehmer Weise. Die wechselnden Ausstellungen bieten einen erfrischenden und dynamischen Kontrast zu der ständigen Kunstsammlung von Schloss Dagstuhl.

Prof. Reinhard Wilhelm, ehemaliger wissenschaftlicher Direktor des Zentrums, fungierte nach seinem Eintritt in den Ruhestand im April 2014 weiterhin als Betreuer der Ausstellungsaktivitäten von Schloss Dagstuhl. Das Zentrum veranstaltet jährlich etwa drei bis vier Kunstausstellungen für jeweils zwei bis drei Monate.

Während lange Jahre einzelne Künstler und Sammler ihre Werke im Zentrum ausstellten, wurden zwischen 2016 und 2018 im Rahmen eines Kooperationsvertrages zwischen der Saarland-Sporttoto GmbH (kurz Saartoto), der Hochschule der Bildenden Künste Saar (kurz HBKsaar) und Schloss Dagstuhl drei Ausstellungen organisiert. Als bedeutender Förderer von Künstlern besitzt Saartoto einen großen Bestand an Kunstwerken..

Im August 2022 wurde für die Dauer von 2 Jahren, die Zusammenarbeit zwischen Schloss Dagstuhl und dem Institut für aktuelle Kunst im Saarland, vertreten durch den Institutsleiter Dr. Andreas Bayer, beschlossen. Die Zusammenarbeit beinhaltet die kuratorische Ausgestaltung von mindestens zwei Ausstellungen/Jahr in Schloss Dagstuhl. Dankenswerterweise unterstützt der *Verein zur Förderung von Schloss Dagstuhl – Leibniz-Zentrum für Informatik e.V.* die Zusammenarbeit.

Pandemiebedingt fand in 2022 nur eine Ausstellung (siehe Fig. 10.1) statt, die nachfolgend beschrieben ist. Die jeweils aktuellen Ausstellungen sind nach Anmeldung auch für die interessierte Öffentlichkeit zugänglich. Informationen und aktuelle Neuigkeiten finden sich auf der Kunst-Webseite[52] von Dagstuhl.

»THINKING STRUCTURES«

Anlässlich der Feier „32 Jahre Informatik in Schloss Dagstuhl“ und der neu vereinbarten Zusammenarbeit mit dem Institut für aktuelle Kunst im Saarland, fand in Schloss Dagstuhl nach zwei Jahren pandemiebedingter Pause wieder eine Kunstausstellung statt. In der Ausstellung wurden

Art exhibitions are regularly organized in the so-called cloister of the new building. The spacious surroundings, excellent lighting, and dramatic day-to-night contrast offer artists a unique exhibition space. Arranged along the corridor walls, the artworks offset the otherwise ascetic nature of the new building. These temporary exhibits offer a fresh and dynamic counterpoint to the center's permanent collection, which can be found scattered throughout the three buildings.

Prof. Reinhard Wilhelm has continued to supervise the Schloss Dagstuhl art exhibitions following his retirement as the center's Scientific Director in April 2014. The center holds approximately three to four art exhibits per year, with each exhibit generally running for two to three months.

For many years Dagstuhl's exhibitions were organized by artists and individual collectors. In the period from 2016 to 2018 three exhibitions had been organized in the course of a cooperation between between Saarland-Sporttoto GmbH (Saartoto for short), Hochschule für Bildende Künste Saar (university of art and design; HBKsaar for short), and Schloss Dagstuhl. Being a major art sponsor, Saartoto is in possession of a substantial art collection.

In August 2022, the cooperation between Schloss Dagstuhl and the Institut für aktuelle Kunst im Saarland (institute for contemporary art), represented by the Institute's director Dr. Andreas Bayer, was agreed for a period of two years. The cooperation includes the curatorial organization of at least two exhibitions/year at Schloss Dagstuhl. The association *Verein zur Förderung von Schloss Dagstuhl – Leibniz-Zentrum für Informatik e.V.* ("Friends of Dagstuhl" for short) gratefully supports the cooperation.

Due to the Covid-19 pandemic, there was only one exhibition hosted by Schloss Dagstuhl in 2022 (cf. Fig. 10.1). Current exhibitions are open to the interested public upon request. For information and current news, please visit Dagstuhl's art webpage[52].

»THINKING STRUCTURES«

On the occasion of the celebration of "32 years of computer science at Schloss Dagstuhl" and the newly agreed cooperation with the Institut für aktuelle Kunst im Saarland, an art exhibition was hosted again at Schloss Dagstuhl after a two-year break due to the pandemic. The exhi-

[52] https://www.dagstuhl.de/de/guests/art

»THINKING STRUCTURES«
Works from Jo Enzweiler, Monika von Boch, and Vera Molnár | June 22 to September 7, 2022

Fig. 10.1
Art exhibition in 2022.

Arbeiten von Monika von Boch (1915–1993), Jo Enzweiler (*1934) und Vera Molnár (*1924) gezeigt. Sie sind wichtige Vertreter eines reflektiert experimentellen Handelns, das die Gestaltungsmöglichkeiten der Bildenden Kunst forschend untersucht.

Monika von Boch visualisiert in ihren Arbeiten elementare Strukturen, die sie mit Fotografie und Manipulation mittels verschiedener fotografischer Techniken darstellt. Jo Enzweiler, ein Vertreter der konkreten Kunst, steht für eine gegenstandsfreie Gestaltung, oft basierend auf rational-mathematischen Gesetzmäßigkeiten. Vera Molnár ist eine Pionierin der digitalen Kunst. Ab 1968 entstanden ihre ersten computergenerierten künstlerischen Arbeiten mit Formen, Linien und Farbwerten. Alle Werke stammen aus der Artothek Saar sowie aus dem Privatbesitz von Jo Enzweiler.

bition featured works by Monika von Boch (1915–1993), Jo Enzweiler (*1934) and Vera Molnár (*1924). They are important representatives of a reflected experimental approach that explores the design possibilities of visual art in an exploratory manner.

Monika von Boch visualizes elementary structures in her works, which she depicts with photography and manipulation using various photographical techniques. Jo Enzweiler, a representative of concrete art, stands for an object-free design, often based on rational-mathematical laws. Vera Molnár is a pioneer of digital art. Starting in 1968, she created her first computer-generated artistic works with shapes, lines and color values. All works come from the Artothek Saar as well as from the private collection of Jo Enzweiler.

Dagstuhls permanente Kunstausstellung

10.2

Dagstuhl's Permanent Art Exhibition

Die von Gästen immer wieder positiv hervorgehobene Kunstsammlung geht auf den Gründungsdirektor Professor Wilhelm zurück. Seine Idee war es, den 1995 neueröffneten Speisesaal und den etwa ein Jahr älteren Neubau durch Kunstwerke zu beleben. Dazu startete er die oben beschrieben Kunstaustellungen. Unter Mitwirkung der Künstler wird aus ausgewählten Ausstellungen ein Werk ausgewählt, für das dann Spender gesucht werden. In den letzten 25 Jahren kamen so ungefähr 180 Kunstwerke zusammen. Auch durch diese Initiative angeregt und verstärkt erhielt Dagstuhl in den vergangenen Jahren weitere Spenden von Künstlern und Mäzenen.

In 2022 spendete so die Künstlerin und Bildhauerin Maria Krause drei Objekte. Die Teilnehmerin Hsiang-Yun Wu spendete ihr Werk „The World of Scientists“, das den ersten Platz auf dem enquote24-Hour Science Map Contest 2021 errang.

Die meisten Arbeiten kommen in den Räumen des Zentrums in Wadern, einige auch in der Geschäftsstelle in Saarbrücken, sehr gut zur Geltung.

The art collection, continually praised by guests, was initiated by Founding Director Professor Wilhelm. It was his idea to use works of art in order to enliven the new building as well as the dining room opened in 1994 and 1995, respectively. To this end, Professor Wilhelm launched the exhibitions described above. Assisted by the artists, an artwork from selected exhibitions is chosen and donors are drummed up. Thus, approximately 180 works of art have been acquired over the last 25 years. Additionally, this initiative has increasingly encouraged artists and patrons to make donations.

In 2022 artist and sculptress Maria Krause donated three objects. Participant Hsiang-Yun Wu donated her work "The World of Scientists," which won first place in the 2021 "24-Hour Science Map Contest."

Most works of art are shown in the center's rooms in Wadern, some are placed in the office in Saarbrücken.

11 Struktur der Gesellschaft
Structure of the Company

Gründung und Gesellschafter 11.1 *Formation and Shareholders*

Schloss Dagstuhl ist als eine gemeinnützige GmbH mit elf Gesellschaftern (siehe Fig. 11.1) organisiert. Dies sind einerseits die vier Gesellschafter, die Schloss Dagstuhl gegründet haben, nämlich die Gesellschaft für Informatik e. V. (GI), die Universität des Saarlandes, die Technische Universität Kaiserslautern und das Karlsruher Institut für Technologie (KIT). Als vier weitere Gesellschafter wurden 1994 die Technische Universität Darmstadt, die Johann Wolfgang Goethe-Universität Frankfurt am Main, die Universität Stuttgart und die Universität Trier aufgenommen. Drei international renommierte Forschungsinstitute, das Institut National de Recherche en Informatique et en Automatique (INRIA, Frankreich), das Centrum Wiskunde & Informatica (CWI, Niederlande) und die Max-Planck-Gesellschaft (MPG, Deutschland) wurden 2005/2006 als weitere Gesellschafter aufgenommen.

Aufgrund eines Beschlusses der Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (heute Gemeinsame Wissenschaftskonferenz) wurde das Zentrum mit Wirkung zum 1. Januar 2006 als Serviceeinrichtung für die Forschung in die gemeinsame Forschungsförderung von Bund und Ländern aufgenommen. Es ist seit 2005 Mitglied der Leibniz-Gemeinschaft. Entsprechend wurde 2008 der Name des Zentrums von vormals „Internationales Begegnungs- und Forschungszentrum für Informatik“ in „Schloss Dagstuhl – Leibniz-Zentrum für Informatik“ geändert.

Schloss Dagstuhl wurde im Juli 2009 erstmals durch die Leibniz-Gemeinschaft evaluiert. Die Stellungnahme der Evaluierungs-Kommission vom März 2010 war sehr positiv: Schloss Dagstuhl widme sich mit herausragendem Erfolg seiner Aufgabe, die internationale Informatikforschung mit einem Seminarzentrum für wissenschaftliche Veranstaltungen zu unterstützen. Schloss Dagstuhl wurde 2016 erneut mit hervorragendem Ergebnis evaluiert. In der Stellungnahme des Senats der Leibniz-Gemeinschaft wurde das Veranstaltungsprogramm und die Beteiligung an der Literaturdatenbank dblp als „exzellent“ bewertet, während der Bereich Open Access (Publishing) als „sehr gut“ bewertet wurde.

Schloss Dagstuhl is operated as a non-profit organization by eleven associates (cf. Fig. 11.1), including its four founding associates: the Gesellschaft für Informatik e. V. (GI),[53] the Universität des Saarlandes, the Technische Universität Kaiserslautern, and the Karlsruher Institut für Technologie (KIT). In 1994, the organization was extended to include four new associates: the Technische Universität Darmstadt, the Johann Wolfgang Goethe-Universität Frankfurt am Main, the Universität Stuttgart, and the Universität Trier. Finally, in 2005 and 2006, three internationally renowned research institutes joined the association: the Institut National de Recherche en Informatique et en Automatique (INRIA, France), the Centrum Wiskunde & Informatica (CWI, Netherlands), and the Max-Planck-Gesellschaft (MPG, Germany).

By resolution of the Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung[54] (today Joint Science Conference) the center has been classified as a research service institution for joint funding by the German federal and state governments since January 2006. Since 2005, Schloss Dagstuhl has been a member of the Leibniz Association and changed its name accordingly from “Internationales Begegnungs- und Forschungszentrum für Informatik”[55] to “Schloss Dagstuhl – Leibniz-Zentrum für Informatik”[56] in 2008.

In July 2009, Schloss Dagstuhl was evaluated for the first time by the Leibniz Association. The March 2010 findings of the evaluation commission were very positive, and established that the center has shown outstanding commitment to its designated task of supporting the international computer science research community by providing a seminar center for academic events. In 2016, Schloss Dagstuhl was evaluated again, with excellent results. In the Leibniz Association Senate report, the seminar program and the cooperation with the computer science bibliography dblp were rated as “excellent” whereas the Open Access Publishing was rated “very good.”

Organe der Gesellschaft 11.2 *Dagstuhl Organs*

Die drei Organe von Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, die stellvertretend für die Gesellschaft als juristische Person handeln, sind die folgenden:

- Gesellschafterversammlung
- Aufsichtsrat
- Geschäftsführung

Details zu den Organen sind den folgenden Abschnitten zu entnehmen.

The three organs of Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, which act for the company as a legal entity, are the following:

- Shareholders’ Meeting
- Supervisory Board
- Management

Detailed information is given in the sections below.

[53] engl.: German Informatics Society
[54] engl.: Federal/State Government Commission for Educational Planning and Research Promotion
[55] engl.: International Conference and Research Center for Computer Science
[56] engl.: Schloss Dagstuhl – Leibniz Center for Informatics

Die Gesellschafterversammlung

Die Gesellschafter beschließen über alle Änderungen an der Gesellschaft, insbesondere über die Aufnahme weiterer Gesellschafter, über die Änderung des Gesellschaftsvertrags und über ihre Auflösung. Die Gesellschafter bestätigen unter anderem auch die von Gesellschaftern neu entsandten Mitglieder in den Aufsichtsrat sowie die Berufung und Abberufung der Geschäftsführer. Derzeit haben anteilig nach der Höhe der Geschäftsanteile alle Gesellschafter die gleiche Anzahl von Stimmen, außer der Gesellschaft für Informatik, die die dreifache Anzahl besitzt. Beschlüsse werden entweder in der mindestens einmal jährlichen stattfindenden Gesellschafterversammlung gefasst oder durch schriftliche Stimmabgabe.

Der Aufsichtsrat

Der Aufsichtsrat ist verantwortlich dafür, dass die Geschäftsführung die Ziele der Gesellschaft rechtmäßig, zweckmäßig und wirtschaftlich sinnvoll erfüllt. Er wirkt in allen wesentlichen Angelegenheiten der Gesellschaft betreffend Forschung und Finanzplanung mit.

Die 12 Mitglieder des Aufsichtsrats (siehe Fig. 11.2) setzen sich aus vier Repräsentanten der Gesellschaft für Informatik, je einem Vertreter der drei Gründungsuniversitäten, zwei Vertretern der später hinzugekommenen vier Universitäten und je einem Vertreter des Bundes und der beiden Bundesländer Saarland und Rheinland-Pfalz, in denen Schloss Dagstuhl formal seinen Sitz hat, zusammen. Die reguläre Amtszeit der Aufsichtsratmitglieder beträgt mindestens vier volle, abgeschlossene Geschäftsjahre und endet mit der Entlastung für das vierte Geschäftsjahr. Die Vertreter der Universitäten in Darmstadt und Stuttgart wechseln im Allgemeinen Amtszeit für Amtszeit mit denen der Universitäten in Frankfurt und Trier ab.

Der Aufsichtsrat entscheidet über die Berufung und Abberufung der Geschäftsführer sowie der Mitglieder des Wissenschaftlichen Direktoriums, des Wissenschaftlichen Beirates und des Kuratoriums. Alle Beschlüsse, die die Finanzen oder das Vermögen der Firma betreffen, benötigen seine Zustimmung. Beschlüsse von forschungspolitischer Bedeutung und Beschlüsse mit erheblichen finanziellen Auswirkungen können nicht gegen die Stimmen der Vertreter des Bundes und der beiden Sitzländer gefasst werden. Der Aufsichtsrat entscheidet zudem über die Erteilung einer Prokura.

Die Geschäftsführung

Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH hat zwei Geschäftsführer (siehe Fig. 11.3), die gemeinsam die Gesellschaft vertreten. Die Geschäftsführung besteht aus dem *Wissenschaftlichen Direktor* und dem *Technisch-administrativen Geschäftsführer*.

Der Wissenschaftliche Direktor ist verantwortlich für die wissenschaftlich-fachliche Zielsetzung und die Programmgestaltung, und ist zudem Mitglied und Vorsitzender des Wissenschaftlichen Direktoriums. Seit Mai 2014 ist Prof. Raimund Seidel, Ph. D., der wissenschaftliche Direktor von Schloss Dagstuhl.

Shareholders' Meeting

All changes to the company, in particular the inclusion of new associates, the revision of the Shareholders' agreement, and the dissolution of the company, are decided by the shareholders. Shareholders also confirm new members forwarded by them to the Supervisory Board and the appointment or recall of the managing directors. In accordance with their shares, all shareholders currently have the same number of votes except the Gesellschaft für Informatik, which has three times the number of votes of the other shareholders in proportion to its larger number of shares. Decisions are made in shareholders' meetings which take place at least once a year, or via a written vote.

Supervisory Board

The Supervisory Board is responsible for ensuring that the management complies with the center's objectives in a legally and economically meaningful manner. The board is involved in all essential matters with regard to research and financial planning.

The 12-member board (see Fig. 11.2) is composed of four representatives of the Gesellschaft für Informatik, one representative from each of the three founding universities, two representatives of the four universities that subsequently joined, and one representative from each of the German federal government and the two host state governments of Saarland and Rhineland-Palatinate. The Supervisory Board members typically hold office for at least four full fiscal years. The term of office ends with the approval for the fourth fiscal year. In general, representatives of the universities in Darmstadt and Stuttgart and of the universities in Frankfurt and Trier rotate after each term of office.

The Supervisory Board formally appoints and recalls the managing directors and members of the Scientific Directorate, Scientific Advisory Board, and Industrial Curatory Board. Furthermore, all decisions regarding financial issues and company assets must be approved by the Supervisory Board. Consent cannot be given against the votes of the represented (federal) state governments if the matter affects political issues in the area of science or has considerable financial weight. The Supervisory Board also holds decision power with respect to the granting of power of attorney.

Management

Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH has two managing directors (see Fig. 11.3) who jointly represent the company. These are the *Scientific Director* and the *Technical Administrative Director*.

The Scientific Director is in charge of drafting the company's scientific goals and program planning, and is also a member and the chairperson of the Scientific Directorate. Since May 2014, Prof. Raimund Seidel, Ph.D., is the Scientific Director of Schloss Dagstuhl.

The Supervisory Board appoints the Scientific Director on basis of the recommendation of a selection committee

Der Wissenschaftliche Direktor wird dem Aufsichtsrat von einer Findungskommission zur Berufung vorgeschlagen. Dieser Findungskommission gehören mindestens der Vorsitzende des Aufsichtsrats und der Vorsitzende des Wissenschaftlichen Beirats an. Die Amtszeit des Wissenschaftlichen Direktors beträgt fünf Jahre.

Die technischen und administrativen Aufgaben werden vom Technisch-administrativen Geschäftsführer wahrgenommen. Seit Juli 2014 hat Heike Meißner diese Position inne.

consisting of at least the chairperson of the Supervisory Board and the chairperson of the Scientific Advisory Board. The term of office of the Scientific Director is five years.

The Technical Administrative Director is responsible for technical and administrative tasks. Since July 2014, Heike Meißner holds this position.

Gremien der Gesellschaft 11.3 *Dagstuhl Bodies*

Die Organe von Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH werden durch drei Gremien unterstützt. Es sind die folgenden:

- Wissenschaftliches Direktorium
- Wissenschaftlicher Beirat
- Kuratorium

Details zu den Gremien werden in den folgenden Abschnitten ausgeführt.

The organs of Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH are supported by the following bodies:

- Scientific Directorate
- Scientific Advisory Board
- Industrial Curatory Board

Detailed information about these boards can be found in the sections below.

Das Wissenschaftliche Direktorium

Das Wissenschaftliche Direktorium (siehe Fig. 11.4) ist für die Realisierung des Gesellschaftszwecks in fachlich-wissenschaftlicher Hinsicht verantwortlich. Es hat das Forschungs- und Veranstaltungsprogramm der Gesellschaft festzulegen, seine fachlich-wissenschaftliche Qualität zu sichern und seine Durchführung zu überwachen. Als wesentlicher Bestandteil dieser Aufgabe werden die Anträge auf Dagstuhl-Seminare und Dagstuhl-Perspektiven-Workshops von Mitgliedern des Wissenschaftlichen Direktoriums begutachtet. Auf den zweimal im Jahr stattfindenden Direktoriumssitzungen werden die Anträge diskutiert und es wird über ihre Annahme entschieden.

Der Wissenschaftliche Direktor gehört dem Wissenschaftlichen Direktorium an. Er empfiehlt dem Aufsichtsrat die Größe des Direktoriums. Neben den Gesellschaftern können das bestehende Wissenschaftliche Direktorium sowie der Beirat Kandidaten für das Wissenschaftliche Direktorium benennen. Die Auswahl der Kandidaten, die dem Aufsichtsrat zur Ernennung vorgeschlagen werden, obliegt dem Beirat zusammen mit dem Wissenschaftlichen Direktor.

Die Amtszeit der Mitglieder des Wissenschaftlichen Direktoriums – mit Ausnahme der des Wissenschaftlichen Direktors – beträgt drei Jahre. Sie beginnt am 1. November des Jahres ihrer Berufung und endet drei Jahre später am 31. Oktober. Wiederberufung ist möglich.

Scientific Directorate

The Scientific Directorate (see Fig. 11.4) is responsible for carrying out the company objectives from a technical and scientific point of view. It must determine the research and event program, ensure its technical and scientific quality, and monitor its execution. As a main task in support of this objective, members of the Scientific Directorate review proposals for Dagstuhl Seminars and Dagstuhl Perspectives Workshops. In its biannual directorate meetings, the Scientific Directorate discusses the proposals and decides which of them to accept or reject.

The Scientific Director is a member of the Scientific Directorate. He recommends to the Supervisory Board the number of Scientific Directorate members. Candidates for the Scientific Directorate may be suggested not only by the shareholders, but also by the Scientific Directorate and the Scientific Advisory Board. The selection of candidates, which are recommended to the Supervisory Board for appointment, is carried out by the Scientific Advisory Board together with the Scientific Director.

The term of office of Scientific Directorate members is three years – with the exception of the Scientific Director. It begins on November 1 of the year of appointment and ends three years later on October 31. Reelections are possible.

Der Wissenschaftliche Beirat

Die Aufgaben des Wissenschaftlichen Beirats (siehe Fig. 11.5) werden nicht nur durch den Gesellschaftsvertrag festgelegt, sondern auch durch die Empfehlungen der Leibniz-Gemeinschaft. Im Sinne dieser wirkt der Wissenschaftliche Beirat auf zwei Wegen bei der Qualitätssicherung mit. Zum einen berät er die Leitung in Fragen der

Scientific Advisory Board

The tasks of the Scientific Advisory Board (see Fig. 11.5) are not only defined by the Shareholders' Agreement, but also by the recommendations of the Leibniz Association. The latter stipulates two different ways in which the Scientific Advisory Board is involved in quality assurance. On the one hand, the board offers advice to the

Forschungs- und Entwicklungsplanung, nimmt Stellung zu den Programmbudgets und gibt Empfehlungen zum Ressourceneinsatz. Er unterstützt weiterhin den Aufsichtsrat bei wichtigen Entscheidungen zur Weiterentwicklung von Schloss Dagstuhl und bei der Gewinnung von Leitungspersonal. Zum anderen führt der Wissenschaftliche Beirat mindestens einmal zwischen je zwei Evaluierungen durch den Senatsausschuss Evaluierung (SAE) der Leibniz-Gemeinschaft ein Audit durch, bei dem die gesamte Einrichtung begutachtet wird. Ein Bericht über das Audit wird der Leitung, dem Aufsichtsrat und dem Senatsausschuss vorgelegt.

Der Wissenschaftliche Beirat sollte aus sechs bis zwölf international angesehenen, im Berufsleben stehenden Wissenschaftlern aus dem In- und Ausland bestehen. Die Amtszeit der Mitglieder beträgt vier Jahre und beginnt am 1. Januar des auf ihrer Berufung folgenden Jahres und endet vier Jahre später am 31. Dezember. Eine einmalige Wiederberufung ist möglich. Der Beirat wählt aus seiner Mitte einen Vorsitzenden. Der Wissenschaftliche Beirat tagt einmal im Jahr. Mitglieder des Beirats werden vom Aufsichtsrat auf Vorschlag des Beirats ernannt.

Das Kuratorium

Das Kuratorium (siehe Fig. 11.6) erfüllt eine Transmissionsfunktion zwischen Schloss Dagstuhl und den Forschungsabteilungen und Entwicklungslaboren der Industrie. Es hat die Aufgabe, die Akzeptanz des Zentrums in Verwaltung, Industrie und Wirtschaft abzusichern und als Förderungsorganisation die wirtschaftliche Basis des Zentrums zu verbreitern. Mitglieder des Kuratoriums werden vom Aufsichtsrat ernannt.

Nach seiner Geschäftsordnung hat das Kuratorium mindestens fünf Mitglieder, deren Amtszeit vier Jahre beträgt. Sie beginnt am 1. Januar des auf ihrer Berufung folgenden Jahres und endet vier Jahre später am 31. Dezember. Eine einmalige Wiederberufung ist möglich. Die Mitglieder des Kuratoriums unterstützen das Zentrum dabei, aktuelle Themen zu identifizieren und dazu geeignete zugkräftige Organisatoren aus der Industrie zu gewinnen. Sie werden ebenso gebeten, geeignete Personen aus der Industrie als Teilnehmer von Dagstuhl-Seminaren und Dagstuhl-Perspektiven-Workshops zu benennen. Das industrielle Kuratorium tagt einmal im Jahr zusammen mit dem Wissenschaftlichen Beirat.

management with regard to research as well as development planning and issues comments on the program budget draft, making recommendations on the use of resources. It also assists the Supervisory Board in the making of important decisions with regard to the future development of the institute as well as the acquisition of management staff. On the other hand, it carries out an audit of the entire institute between two evaluations by the Senatsausschuss Evaluierung (SAE, Senate Committee Evaluation) of the Leibniz Association. A report on this audit is sent to the management, the Supervisory Board, and the SAE.

The Scientific Advisory Board should consist of six to twelve internationally reputable, well established scientists and academics from Germany and abroad. The term of office for members is four years and can be prolonged once. It begins on January 1 of the year after the appointment and ends four years later on December 31. The Scientific Advisory Board members elect a chairperson from their midst. The board convenes once a year. Members are appointed by the Supervisory Board in accordance with the suggestions of the Scientific Advisory Board.

Industrial Curatory Board

The Industrial Curatory Board (see Fig. 11.6) performs a transmissional function between the center and the industrial R&D departments and laboratories. Its role is to secure acceptance of Schloss Dagstuhl within the business, industry and administrative communities, and as a promotional organization to broaden the economic basis of the center. Board members are appointed by the Supervisory Board.

According to its rules of procedure, the Industrial Curatory Board consists of at least five members whose term of office is four years. It begins on January 1 of the year after the appointment and ends four years later on December 31. A one-off reappointment for a second term is possible. The board members help the center to identify current R&D topics for seminars and locate attractive organizers in industry. The Industrial Curatory Board is regularly called upon to propose suitable participants for Dagstuhl Seminars and Dagstuhl Perspectives Workshops known to it from its activities. It convenes once a year together with the Scientific Advisory Board.

Gesellschafter | Associates
Centrum Wiskunde & Informatica (CWI), The Netherlands
Gesellschaft für Informatik e. V., Germany
Institut National de Recherche en Informatique et en Automatique (INRIA), France
Johann Wolfgang Goethe-Universität Frankfurt am Main, Germany
Karlsruher Institut für Technologie (KIT), Germany
Max-Planck-Gesellschaft zur Förderung der Wissenschaften e. V., Berlin, Germany
Technische Universität Darmstadt, Germany
Technische Universität Kaiserslautern, Germany
Universität des Saarlandes, Germany
Universität Stuttgart, Germany
Universität Trier, Germany

Fig. 11.1
Associates.

Aufsichtsrat | Supervisory Board
Dr. Marc Brüser Ministerium für Wissenschaft, Weiterbildung und Kultur, Mainz, Germany | Representative of Rhineland-Palatinate state
Prof. Dr. Hannes Hartenstein Karlsruher Institut für Technologie, Germany | Representative of Karlsruher Institut für Technologie
Prof. Dr.-Ing. Dr. h. c. Stefan Jähnichen Technische Universität Berlin, Germany | Representative of Gesellschaft für Informatik e. V. | Chairman of the Supervisory Board
Prof. Dr. Volker Lindenstruth Johann Wolfgang Goethe-Universität Frankfurt am Main, Germany | Representative of Johann Wolfgang Goethe-Universität Frankfurt am Main
Dr. Svenja Marx Bundesministerium für Bildung und Forschung, Bonn, Germany | Representative of the German federal government
Christian Mees Staatskanzlei des Saarlandes, Saarbrücken, Germany | Representative of the Saarland | *tenure started in October 2022*
Prof. Dr. Arnd Poetzsch-Heffter Technische Universität Kaiserslautern, Germany | Representative of Technische Universität Kaiserslautern
Christine Regitz SAP SE | Representative of Gesellschaft für Informatik e. V. | *tenure started in May 2022*
Dr. Susanne Reichrath Staatskanzlei des Saarlandes, Saarbrücken, Germany | Representative of the Saarland | *tenure ended in May 2022*
Prof. Dr. Ralph Schenkel Universität Trier, Germany | Representative of Universität Trier
Prof. Dr. Manfred J. Schmitt Universität des Saarlandes, Saarbrücken, Germany | Representative of Universität des Saarlandes
Prof. em. Dr.-Ing. Dr.-Ing. h. c. Roland Vollmar Karlsruher Institut für Technologie, Germany | Representative of Gesellschaft für Informatik e. V.
Cornelia Winter Gesellschaft für Informatik e. V., Bonn, Germany | Representative of Gesellschaft für Informatik e. V

Fig. 11.2
Supervisory Board members.

Geschäftsführung | Management
Heike Meißner (Technisch-administrative Geschäftsführerin | Technical Administrative Director) Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Wadern, Germany
Prof. Raimund Seidel, Ph. D. (Wissenschaftlicher Direktor | Scientific Director) Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Wadern and Universität des Saarlandes, Saarbrücken, Germany

Fig. 11.3
Management.

Wissenschaftliches Direktorium | Scientific Directorate
Prof. Dr. Elisabeth André Universität Augsburg, Germany
Prof. Dr.-Ing. Franz Baader TU Dresden, Germany
Prof. Dr. Daniel Cremers Technische Universität München, Germany
Goetz Graefe, Ph. D. Google Inc., Madison, United States of America
Prof. Dr. Reiner Hähnle TU Darmstadt, Germany
Prof. Dr. Barbara Hammer Universität Bielefeld, Germany
Prof. Dr. Lynda Hardman Centrum Wiskunde & Informatica (CWI), Amsterdam and University of Utrecht, The Netherlands
Prof. Dr.-Ing. Oliver Kohlbacher Eberhard Karls Universität Tübingen, Germany
Dr. Steve Kremer Institut National de Recherche en Informatique et en Automatique (INRIA), Nancy – Grand Est, France
Prof. Rupak Majumdar, Ph. D. Max Planck Institute for Software Systems, Kaiserslautern, Germany
Prof. Dr.-Ing. Heiko Mantel TU Darmstadt, Germany
Prof. Dr. Albrecht Schmidt Ludwig-Maximilians Universität München, Germany
Prof. Dr.-Ing. Wolfgang Schröder-Preikschat Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany
Prof. Raimund Seidel, Ph. D. Universität des Saarlandes, Saarbrücken, Germany
Prof. Dr. Heike Wehrheim Universität Paderborn, Germany
Prof. Dr. Verena Wolf Universität des Saarlandes, Saarbrücken, Germany
Prof. Dr. Martina Zitterbart Karlsruher Institut für Technologie, Germany

Fig. 11.4
Scientific Directorate.

Wissenschaftlicher Beirat | Scientific Advisory Board
Prof. Dr. Christel Baier Technische Universität Dresden, Germany | Chair of the Scientific Advisory Board
Prof. Dr. Anja Feldmann Max-Planck-Institut für Informatik, Saarbrücken, Germany
Prof. Dr. Ir. Dr. h. c. (AAU) Joost-Pieter Katoen RWTH Aachen, Germany
Prof. Dr. Laurent Romary Centre Marc Bloch, Berlin, Germany and INRIA, Paris, France
Prof. Alistair Sinclair, Ph. D. University of California, Berkeley, United States of America | *tenure started in January 2022*
Prof. Dr. Lothar Thiele ETH Zürich, Switzerland | *tenure started in January 2022*

Fig. 11.5
Scientific Advisory Board.

Kuratorium | Industrial Curatory Board
Dr. Tim Harris Microsoft Research, Cambridge, United Kingdom
Dr. Jaroslaw Kutylowski DeepL, Köln, Germany | *tenure started in January 2022*
Dr.-Ing. Christof Leng Google | *tenure will start in January 2023*
Christine Regitz SAP SE | *resigned in May 2022*
Dr.-Ing. Andreas Wierse SICOS BW GmbH, Stuttgart, Germany
Dr. Thomas Ziegert SAP SE, Darmstadt, Germany

Fig. 11.6
Industrial Curatory Board.

12 Förderverein „Freunde von Dagstuhl“
Association “Friends of Dagstuhl”

Förderverein „Freunde von Dagstuhl“

Association “Friends of Dagstuhl”

Holger Hermanns (Universität des Saarlandes, Germany)
Erich Reindel (Universität des Saarlandes, Germany)

Seit Mitte 2014 gibt es den Verein zur Förderung von Schloss Dagstuhl – Leibniz-Zentrum für Informatik e.V. Der sehr technische und holprig klingende Name spiegelt dabei exakt den Vereinszweck wider: die Förderung von Wissenschaft und Forschung im Leibniz-Zentrum für Informatik in Schloss Dagstuhl. Für die Webpräsenz wurde allerdings ein wesentlich geschmeidigerer Name gewählt: „Friends of Dagstuhl“ (http://www.friends-of-dagstuhl.de).

Der Verein ist darauf ausgerichtet, finanzielle Mittel zur erfolgreichen Umsetzung des Vereinszwecks zu beschaffen und bereitzustellen sowie die ihm zu diesem Zweck anvertrauten Mittel treuhänderisch zu verwalten. Die Stiftung Informatikzentrum Schloss Dagstuhl wurde daher auch als nicht rechtsfähige Stiftung in den Verein überführt. Seit Ende 2014 vertreten nun die Freunde von Dagstuhl die Stiftung im Rechts- und Geschäftsverkehr und verwalten das Stiftungsvermögen unter der strategischen Aufsicht eines Stiftungsrates (siehe Fig. 12.1). Der Verein wird von einem Vorstand (siehe Fig. 12.2 und Fig. 12.3) geleitet.

Wie bereits im Jahresbericht 2019 dargestellt, ist es wegen einer großzügigen Zuwendung aus dem Nachlass eines verstorbenen Freundes von Schloss Dagstuhl möglich, die Anbindung von Schloss Dagstuhl über den ÖPNV an den Bahnhof Türkismühle deutlich zu erleichtern. Da allerdings während der Covid-19-Pandemie auch die Tagungen in Schloss Dagstuhl zum Erliegen kamen, wurde dieser Plan bisher nicht umgesetzt. Noch in 2023 soll aber endlich ein vom Verein finanzierter Shuttle-Service an den An- und Abreisetagen etabliert werden.

Weiterhin hat der Verein die Stiftung eines *Dagstuhl-Awards* auf den Weg gebracht, durch den in Zukunft herausragende Leistungen in der Informatik oder auch herausragende Dagstuhl-Seminare ausgezeichnet werden. Eine neu gegründete Kommission, die das Konzept und Profil eines solchen Awards entwickeln soll, hat bereits die Arbeit aufgenommen.

Weitere Informationen zum Verein, aber auch Mitgliedschaftsanträge finden Sie unter http://www.friends-of-dagstuhl.de.

Since mid 2014, the registered association for support of Schloss Dagstuhl – Leibniz Center for Informatics (Verein zur Förderung von Schloss Dagstuhl – Leibniz-Zentrum für Informatik e.V.) has existed. This very technical and rather clumsy name nevertheless reflects the precise purpose of the association: the promotion of science and research at the Leibniz Center for Informatics at Schloss Dagstuhl. A significantly smoother name, i.e., “Friends of Dagstuhl”, was chosen for the website (http://www.friends-of-dagstuhl.de).

The association aims to acquire and provide funds for the successful execution of its purposes, as well as to hold the funds entrusted to it. The Dagstuhl Foundation (Stiftung Informatikzentrum Schloss Dagstuhl) was therefore integrated into the association as a dependent foundation. Since late 2014, the Friends of Dagstuhl represent the Dagstuhl Foundation in legal and business transactions and manage the foundation assets under the strategic supervision of a foundation council (see Fig. 12.1). The association is chaired by a board (see Fig. 12.2 and Fig. 12.3).

As already outlined in the annual report for 2019, due to a generous donation from the estate of a deceased friend of Schloss Dagstuhl, it is possible to significantly facilitate the connection of Schloss Dagstuhl via public transport to the train station in Türkismühle. However, as the Covid 19 pandemic also brought the events at Schloss Dagstuhl to a standstill, this plan has not yet been implemented. In 2023, however, a shuttle service financed by the association is finally to be established on the days of arrival and departure.

Furthermore, the association has initiated the foundation of a *Dagstuhl Award*, which in the future will honour outstanding achievements in computer science or Dagstuhl Seminars. A newly founded commission, which is to develop the concept and profile of such an award, has already begun its work.

Further information about the association as well as the membership application form can be found at http://www.friends-of-dagstuhl.de.

Stiftungsrat | Foundation council
Prof. Dr. Holger Hermanns (Vorstandsvorsitzender des Vereins “Friends of Dagstuhl” | First deputy chairperson of the association “Friends of Dagstuhl”) Universität des Saarlandes, Saarbrücken, Germany
Prof. Dr. Dr. h.c. mult. Kurt Mehlhorn Max Planck Institute for Informatics (MPII), Saarbrücken, Germany
Prof. Dr. Dorothea Wagner Karlsruher Institut für Technologie (KIT), Germany

Fig. 12.1
Der Stiftungsrat der Stiftung “Informatik-Zentrum Schloss Dagstuhl”
The council of the foundation “Informatik-Zentrum Schloss Dagstuhl”

Vorstand des Vereins | Chair of the association
Prof. Dr. Holger Hermanns (Vorstandsvorsitzender | First deputy chairperson) Universität des Saarlandes, Saarbrücken, Germany
Angelika Müller-von Brochowski (Schriftführerin | Secretary) Saarbrücken, Germany
Erich Reindel (Schatzmeister | Treasurer) Universität des Saarlandes, Saarbrücken, Germany

Fig. 12.2
Der Vorstand des Vereins “Friends of Dagstuhl”
The chair of the association “Friends of Dagstuhl”

Fig. 12.3
Der Vorstand des Vereins “Friends of Dagstuhl”, v.l.n.r.: Prof. Dr. Holger Hermanns, Angelika Müller-von Brochowski, und Erich Reindel.
The chair of the association “Friends of Dagstuhl”, f.l.t.r.: Prof. Holger Hermanns, Angelika Müller-von Brochowski, and Erich Reindel.

13 Statistiken
Statistics

Statistiken zu Seminaren und Workshops

13.1

Statistics on Seminars and Workshops

In diesem Abschnitt werden statistische Daten zum wissenschaftlichen Programm und der Zusammenstellung der Teilnehmer aufgeführt. Durch die Covid-19 Pandemie bedingt, weichen die veranstaltungs- und teilnehmerbezogenen statistischen Werte der Jahre 2020 und 2021 erheblich von denen der vorhergehenden Jahre bis einschließlich 2019 ab. Details sind in den Abschnitten weiter unten aufgeführt.

Die Diagramme und Tabellen sind dabei wie nachfolgend beschrieben gegliedert.

Antrags-bezogene Daten: Die Anzahl eingereichter Anträge von Dagstuhl Seminaren und Dagstuhl Perspektiven Workshops sowie deren Akzeptanzraten sind in Fig. 13.1 dargestellt. Die sehr hohe Anzahl von Anträgen 2020 kann unter anderem auf die pandemiebedingte auf Juni 2020 verschobene zweite Antragsfrist zurückgeführt werden. In dieser Runde wurde mit 92 Anträge soviele wie nie zuvor gestellt. Die sehr niedrige Anzahl 2021 ist ausschließlich durch eine beschränkte erste Antragsrunde, in der nur wegen der Pandemie abgesagte Seminare neu beantragt werden durften, begründet. In dieser Runde wurden schließlich alle Anträge, die ja schon einmal positiv begutachtet worden waren, genehmigt, was den überdurchschnittlich hohen Anteil der genehmigten Seminare erklärt. Fig. 13.2 zeigt, wie die akzeptierten Seminare und Workshops sich bezüglich Größe und Länge aufgliedern.

Veranstaltungs-bezogene Daten: Fig. 13.3 zeigt die Anzahl der verschiedenen Veranstaltungstypen. Der dramatische Einbruch 2020 und die immer noch geringe Anzahl von Veranstaltungen in 2021 sind ebenso durch die Covid-19 Pandemie bedingt. In 2020 hatte Schloss Dagstuhl wegen der Pandemie ab Mitte März für etwa 5 Monate geschlossen. Aber auch nach Wiedereröffnung sind 2020 fast alle geplanten Veranstaltungen abgesagt worden. Beginnend mit 2021 bietet Schloss Dagstuhl übergangsweise auch hybride Seminare, in denen Teilnehmer sowohl vor Ort sind als auch online über Audio-/Videoübertragung zugeschaltet sind, an. Als Spezialfall sind in den reinen online Seminare gar keine Teilnehmer vor Ort. Alle diese Typen werden in der Tabelle hier mitgezählt. Fig. 13.4 gibt aber einen Überblick, wie viele Veranstaltungen on-site, hybrid oder online stattgefunden haben.

Daten zu der Anzahl der durchgeführten Seminare gegliedert nach der ursprünglich genehmigten Größe und Dauer sind in Fig. 13.5 angegeben. Manche Organisatorenteams von – insbesondere rein online stattfindenden – Seminaren hatten sich entschlossen, die Dauer teilweise erheblich zu verkürzen.

Zum Ausgleich der Einschränkungen von online oder hybrid stattfindenden Seminaren wurden bei diesen auch mehr Teilnehmer als ursprünglich genehmigt zugelassen. Fig. 13.6 zeigt Anzahl und Anteil der eingeladenen Seminarteilnehmer, welche die Einladung

This section provides statistical data about the scientific program and the composition of program participants. Due to the Covid-19 pandemic, the statistical values related to the program and participants of 2020 and 2021 deviate considerably from those of previous years up to and including 2019. Details are provided in the sections below.

Charts and tables are structured as described below.

Proposal-related data: Fig. 13.1 shows the number of submitted proposals for Dagstuhl Seminars and Dagstuhl Perspectives Workshops as well as the respective acceptance rates. The exceptionally large number of proposals in 2020 is due to the fact that the second submission deadline was postponed until June 2020, among other reasons. That proposal round saw 92 submissions, more than ever before. The very low number in 2021 is exclusively due to the fact that the first submission round was limited in that only resubmissions of proposals for seminars that had been cancelled due to the pandemic were allowed. Eventually, all proposals were accepted – they had obviously been accepted before – which explains the above-average share of accepted seminars. Size and duration of accepted seminars and workshops are displayed in Fig. 13.2.

Event-related data: Fig. 13.3 illustrates the number of different event types. The sharp drop in 2020 and the continuously low number of events in 2021 were also caused by the Covid-19 pandemic. From mid-March 2020, Schloss Dagstuhl was closed for approximately 5 months as a consequence of the pandemic. Even after re-opening, almost all scheduled events in 2020 were cancelled. Beginning in 2021 and on an interim basis, Schloss Dagstuhl has been offering hybrid seminars where some participants are on site while others join the seminar online via audio/video transmission. In special cases, there are also online-only seminars where there are no participants on site. The table includes all of these event types. However, Fig. 13.4 provides an overview of how many events took place on site, in a hybrid setting or online, respectively.

Fig. 13.5 illustrates data regarding the number of seminars that took place, categorised by originally approved size and duration. Some teams of seminar organizers – especially in the context of online-only seminars – decided to shorten the duration, in some cases drastically.

In order to compensate for the limitations of hybrid and online seminars, larger numbers of participants than originally approved were allowed to join. Fig. 13.6 shows the number and proportion of invited seminar participants who accepted or declined the invitation. It is evident that the share of invitees who accepted their invitation was significantly higher in 2021. This may

annehmen bzw. ablehnen. Der in 2021 auffallende, deutlich höhere Anteil an Eingeladenen, die zugesagt haben, mag zum einen an der Möglichkeit der Online-Teilnahme eines Teils der angebotenen Seminare liegen, aber genauso auch daran, dass zu manchen Seminaren noch sehr kurzfristig vor dem Seminar Personen eingeladen wurden, deren Zusage vor der Einladung abgesprochen wurde. Auch dies wurde, zum Teil, durch die Möglichkeit der Online-Teilnahme ermöglicht.

Die Verteilung der Annahmerate pro Seminar, also der Quotient aus Teilnehmer und allen Eingeladenen, ist in Fig. 13.7 dargestellt. Durch die pandemiebedingten Besonderheiten des Seminarbetriebs in und den deutlich kleineren Stichproben aus 2020 und 2021 sind die Daten dieser Jahre nicht im Vergleich der anderen Jahre zu interpretieren.

Fig. 13.8 zeigt dagegen, wie viel Prozent der zugesagten Größe (gemessen an der Personenanzahl) tatsächlich von einem Seminar belegt wurde. Auch hier sind die Daten der letzten beiden Jahre nicht mit den vorherigen Daten zu vergleichen. Werte über 100 % bedeuten, dass mehr Personen als ursprünglich genehmigt am Seminar teilgenommen haben. Während dies bei online zugeschalteten Teilnehmer problemlos möglich ist, war die Gesamtanzahl der Teilnehmer im Haus und damit auch für die Veranstaltungen strikter limitiert.

Teilnehmer-bezogene Daten: Grundsätzlich werden hier alle Teilnehmer gezählt, unabhängig davon, ob sie vor Ort oder online teilgenommen haben. Vor allem durch ausfallende Veranstaltungen aber auch durch schlechtere Beteiligung insbesondere an Veranstaltungen die ausschließlich vor Ort stattfanden, sind die Zahlen aus 2020 und 2021 deutlich geringer als die der Vorjahre.

Die Teilnehmerzahlen – abhängig vom Veranstaltungstyp – gibt Fig. 13.9 an. Einen Eindruck, wieviele Teilnehmer vor Ort waren bzw. remote teilgenommen haben, vermittelt Fig. 13.10.

Fig. 13.11 zeigt die Verteilung der Herkunftsländer unserer Gäste.

Umfrage-bezogene Daten: Hier stellen wir ausgewählte Daten unserer fortlaufenden Befragung von Teilnehmern an Dagstuhl-Seminaren und Dagstuhl-Perspektiven-Workshops dar. Ein Überblick über die Ergebnisse der regelmäßigen Gästebefragungen kann Fig. 13.12 entnommen werden. Die Anzahl von früheren Seminarbesuchen kann man Fig. 13.13 entnehmen. Fig. 13.14 gibt Auskunft über die Altersstruktur der Teilnehmer. Während Dagstuhl-Seminare und Dagstuhl-Perspektiven-Workshops sich primär an Forscher aus Universitäten und Forschungseinrichtungen richten, sind auch Anwender und Forscher aus der Industrie stets willkommen. Die Verteilung ihres Anteils ist in Fig. 13.15 gezeigt.

Auslastungs-bezogene Daten: Die Auslastung des Zentrums wird schließlich in Fig. 13.16 an Hand der Übernachtungen und ihrer Verteilung über die einzelnen Wochen getrennt nach Veranstaltungstypen aufgezeigt. Im Gegensatz zu den anderen Statistiken beziehen sich diese Daten ausschließlich auf Gäste vor Ort.

in part be due to the fact that some seminars offered online participation, but may also have been caused by issuing invitations to some seminars at very short notice when the invitees had informally agreed to attend in advance. This was also partly made possible through offering online participation.

The acceptance rate distribution per seminar, i.e., the quotient of the number of participants and the total number of invitees, is illustrated in Fig. 13.7. Due to the particularities regarding the seminars and the significantly smaller samples in 2020 and 2021 caused by the pandemic, the data from those years is not to be interpreted in relation to previous years.

In contrast, Fig. 13.8 visualizes the percentage of the reserved space (in terms of number of people) that was actually used by seminar participants. Again, the data from the last two years cannot be compared to previous data. Values above 100 percent show that more people than originally approved attended the seminar. While this was easy to facilitate with regard to online participants, the total number of participants on site and therefore the maximum number of on-site participants allowed for the respective events was limited more strictly.

Participant-related data: Basically, all participants are counted here, regardless of whether they participated on site or online. The numbers from 2020 and 2021 are significantly lower than the ones in previous years due especially to event cancellations, but also due to lower participation rates, particularly with regard to on-site only events.

Fig. 13.9 shows the number of participants broken down by event type. Fig. 13.10 illustrates how many participants were on site and how many participated remotely.

Fig. 13.11 shows the distribution of our guests' country affiliations.

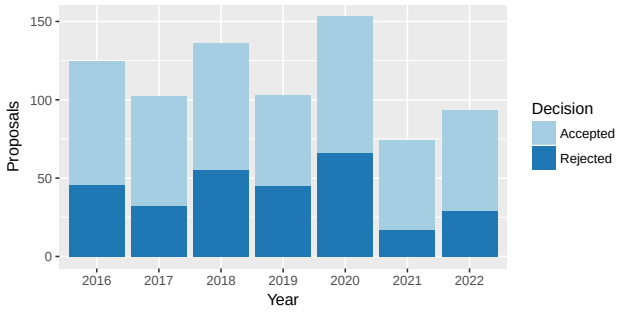
Survey-related data: In this section we present selected data obtained from our ongoing Dagstuhl Seminar and Dagstuhl Perspectives Workshop guest survey project. An overview of the results gained from routine participants surveys for Dagstuhl Seminars and Dagstuhl Perspectives Workshops can be found in Fig. 13.12. Fig. 13.13 displays how often participants have attended seminars in the past. Fig. 13.14 provides data on the seniority of participants. While Dagstuhl Seminars and Dagstuhl Perspectives Workshops mainly target academic researchers, researchers and developers from industry are also always welcome. The distribution of their share is shown in Fig. 13.15.

Utilization-related data: Finally, Fig. 13.16 illustrates utilization on the basis of overnight stays hosted at Schloss Dagstuhl – broken down by event type – as well as their distribution by week. In contrast to the other statistics, this data exclusively refers to guests on site.

Gender-related data: Fig. 13.17 illustrates gender distribution with respect to organizer teams of Dagstuhl Seminars and Dagstuhl Perspectives Workshops. In contrast, Fig. 13.18 shows the proportion of women

Geschlechter-bezogene Daten: Fig. 13.17 enthält Daten zur Geschlechter-Verteilung in der Seminarleitung. Dagegen zeigt Fig. 13.18 die Quote von Frauen bei der Beantragung von Seminaren sowohl bezüglich der Teams als auch bezüglich der gesamten Antragsteller. Die Abbildungen Fig. 13.19 und Fig. 13.20 zeigen insbesondere die Anteile weiblicher Teilnehmer bzw. Einladungen an weibliche Wissenschaftler. Die Verteilung der Rate der weiblichen Teilnehmer in den einzelnen Seminaren wird in Fig. 13.21 aufgezeigt. Diese Statistiken werden von unsere Geldgeber im Rahmen der Frauenförderung gefordert. Die Daten basieren vorwiegend auf der Selbstauskunft von Gästen, dass sie weiblich sind.

in seminar proposals with respect to both the teams and the proposers overall. Fig. 13.19 and Fig. 13.20 mainly illustrate the share of female participants and invitees, respectively. The distribution of the share of women among seminar participants is displayed in Fig. 13.21. Please note that our funding agencies require such statistics as part of their efforts to promote gender equality. The statistics are predominantly based on guests' self-reporting that they are female.

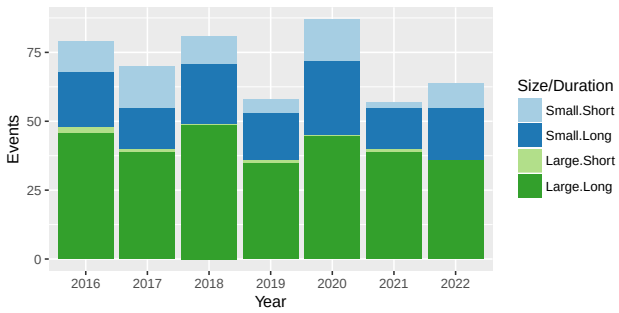


(a) Chart for 2016–2022

Year	Proposals	Accepted		Rejected	
	#	#	%	#	%
2016	125	79	63.2	46	36.8
2017	102	70	68.6	32	31.4
2018	136	81	59.6	55	40.4
2019	103	58	56.3	45	43.7
2020	153	87	56.9	66	43.1
2021	74	57	77.0	17	23.0
2022	93	64	68.8	29	31.2

(b) Detailed numbers for 2016–2022

Fig. 13.1
Proposals and acceptance rates for Dagstuhl Seminars and Dagstuhl Perspectives Workshops.

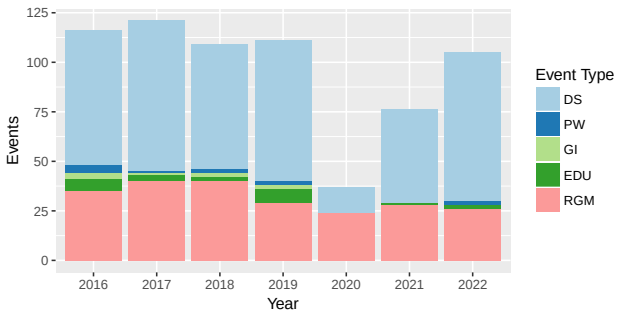


(a) Chart for 2016–2022

Year	30-person seminars		45-person seminars		Total
	3-day	5-day	3-day	5-day	
2016	11	20	2	46	79
2017	15	15	1	39	70
2018	10	22	0	49	81
2019	5	17	1	35	58
2020	15	27	0	45	87
2021	2	15	1	39	57
2022	9	19	0	36	64

(b) Detailed numbers for 2016–2022

Fig. 13.2
Size and duration of Dagstuhl Seminars and Dagstuhl Perspectives Workshops accepted in 2016–2022. Small = 30-person seminar, Large = 45-person seminar, Short = 3-day seminar, Long = 5-day seminar.

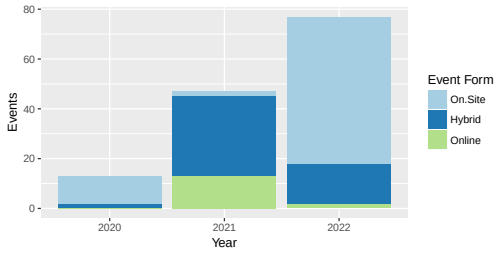


(a) Chart for 2016–2022

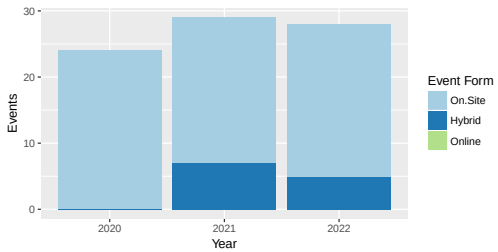
Year	DS	PW	GI	EDU	RGM	Total
2016	68	4	3	6	35	116
2017	76	1	1	3	40	121
2018	63	2	2	2	40	109
2019	71	2	2	7	29	111
2020	13	0	0	0	24	37
2021	47	0	0	1	28	76
2022	75	2	0	2	26	105

(b) Detailed numbers for 2016–2022

Fig. 13.3
Number of all events held at Dagstuhl, by type. DS = Dagstuhl Seminar, PW = Dagstuhl Perspectives Workshop, GI = GI-Dagstuhl Seminar, EDU = educational event, RGM = research group meeting.



(a) Graphical distribution of seminars in group A in 2020–2022 by year and form

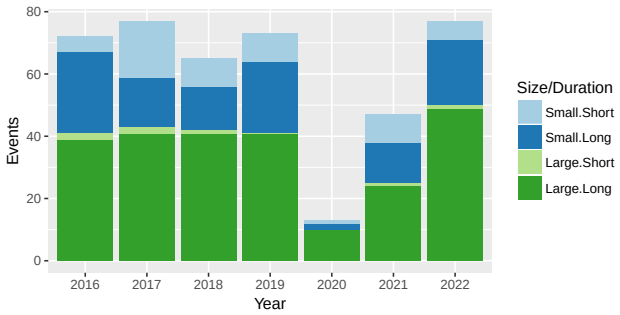


(b) Graphical distribution of events in group B in 2020–2022 by year and event form

	Group A			Group B		
Year	On-Site	Hybrid	Online	On-Site	Hybrid	Online
2020	11	2	0	24	0	0
2021	2	32	13	22	7	0
2022	59	16	2	23	5	0

(c) Detailed numbers for 2020–2022 by event form

Fig. 13.4
Number of all events held at Dagstuhl, by event form and group. Group A = Dagstuhl Seminars and Dagstuhl-Perspectives-Workshops. Group B = all other events (GI-Dagstuhl Seminars, educational events, and research group meetings).

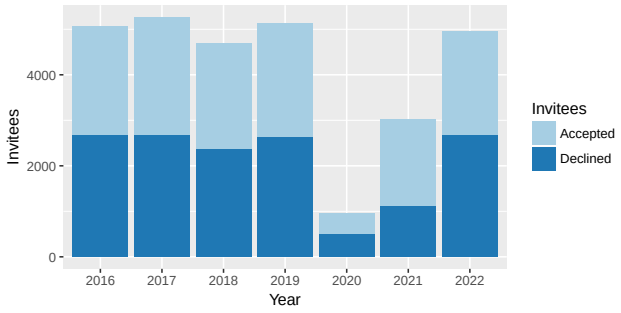


(a) Chart for 2016–2022

Year	30-person seminars		45-person seminars		Total
	3-day	5-day	3-day	5-day	
2016	5	26	2	39	72
2017	18	16	2	41	77
2018	9	14	1	41	65
2019	9	23	0	41	73
2020	1	2	0	10	13
2021	9	13	1	24	47
2022	6	21	1	49	77

(b) Detailed numbers for 2016–2022

Fig. 13.5
Size and duration of Dagstuhl Seminars and Dagstuhl Perspectives Workshops held in 2016–2022. Small = 30-person seminar, Large = 45-person seminar, Short = 3-day seminar, Long = 5-day seminar.

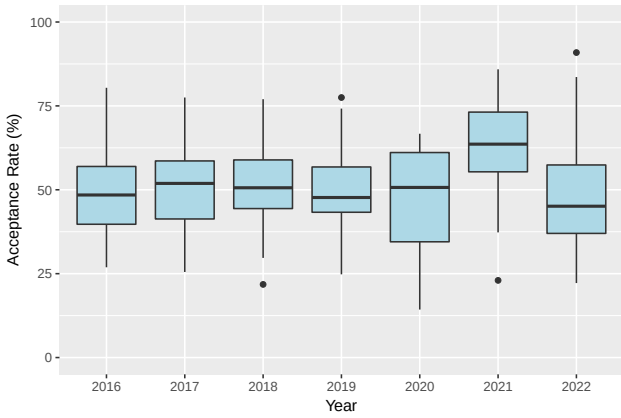


(a) Chart for 2016–2022

Year	Invitees	Accepted		Declined	
	#	#	%	#	%
2016	5060	2393	47.3	2667	52.7
2017	5267	2572	48.8	2695	51.2
2018	4692	2320	49.4	2372	50.6
2019	5143	2498	48.6	2645	51.4
2020	964	442	45.9	522	54.1
2021	3022	1894	62.7	1128	37.3
2022	4950	2278	46.0	2672	54.0

(b) Detailed numbers for 2016–2022

Fig. 13.6
Total number of invitees, accepted and declined invitations for Dagstuhl Seminars and Dagstuhl Perspectives Workshops.

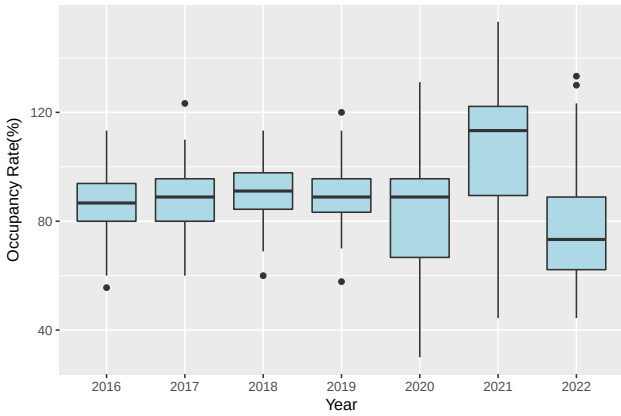


(a) Chart for 2016–2022

Year	Min (%)	Max (%)	Avg (%)	Std (%)
2016	26.9	80.4	48.6	11.2
2017	25.5	77.5	50.3	12.4
2018	21.8	77.0	51.2	12.0
2019	24.8	77.5	49.8	11.4
2020	14.3	66.7	45.8	17.4
2021	23.0	85.9	63.2	12.5
2022	22.2	90.9	47.9	14.1

(b) Detailed numbers for 2016–2022

Fig. 13.7
Acceptance rate distribution per Dagstuhl Seminar or Dagstuhl Perspectives Workshop in 2016–2022. Min = minimal value, Max = maximal value, Avg = average, Std = standard deviation.

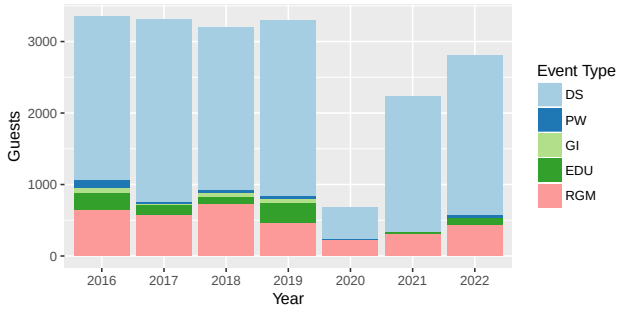


(a) Chart for 2016–2022

Year	Min (%)	Max (%)	Avg (%)	Std (%)
2016	55.6	113.3	86.7	11.8
2017	60.0	123.3	87.3	12.3
2018	60.0	113.3	90.3	10.2
2019	57.8	120.0	89.1	10.7
2020	30.0	131.1	79.7	27.9
2021	44.4	153.3	108.0	24.6
2022	44.4	133.3	76.1	18.7

(b) Detailed numbers for 2016–2022

Fig. 13.8
Occupancy rate distribution per Dagstuhl Seminar or Dagstuhl Perspectives Workshop in 2016–2022. Min = minimal value, Max = maximal value, Avg = average, Std = standard deviation.

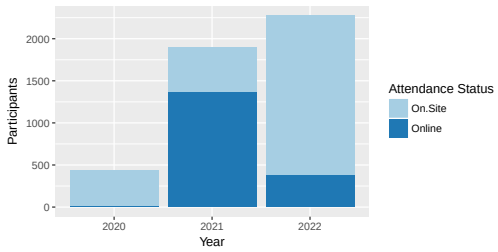


(a) Chart for 2016–2022

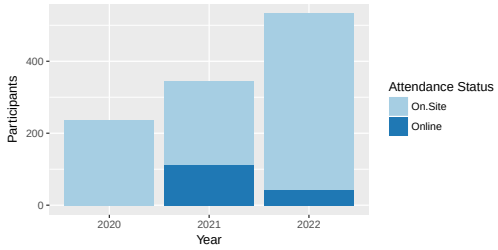
Year	DS		PW		GI		EDU		RGM		Total
	#	%	#	%	#	%	#	%	#	%	#
2016	2280	68.0	113	3.4	78	2.3	232	6.9	650	19.4	3353
2017	2551	77.1	21	0.6	21	0.6	131	4.0	584	17.7	3308
2018	2268	70.8	52	1.6	50	1.6	99	3.1	733	22.9	3202
2019	2450	74.3	48	1.5	50	1.5	282	8.5	469	14.2	3299
2020	442	65.1	0	0.0	0	0.0	0	0.0	237	34.9	679
2021	1894	84.6	0	0.0	0	0.0	31	1.4	314	14.0	2239
2022	2229	79.3	49	1.7	0	0.0	97	3.5	436	15.5	2811

(b) Detailed numbers for 2016–2022

Fig. 13.9
Number of participants. DS = Dagstuhl Seminar, PW = Dagstuhl Perspectives Workshop, GI = GI-Dagstuhl Seminar, EDU = educational event, RGM = research group meeting.



(a) Graphical distribution of participants of seminars in group A in 2020–2022 by year and attendance status



(b) Graphical distribution of participants of events in group B in 2020–2022 by year and attendance status

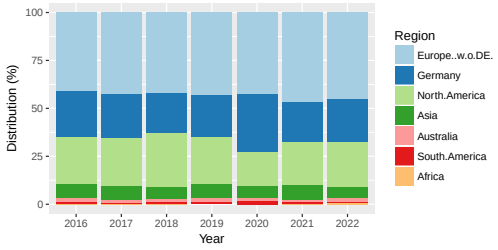
	Group A			Group B		
Year	On-Site	Online	Total	On-Site	Online	Total
2020	419	23	442	237	0	237
2021	529	1365	1894	233	112	345
2022	1895	383	2278	491	42	533

(c) Detailed numbers for 2020–2022 by attendance status

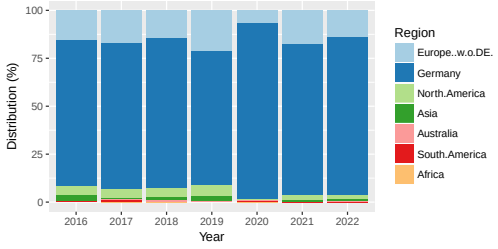
Fig. 13.10
Number of participants by attendance status and group. Group A = Dagstuhl Seminars and Dagstuhl-Perspectives-Workshops. Group B = all other events (GI-Dagstuhl Seminars, educational events, and research group meetings).

Country	A	B	Total
Germany	510	437	947
United States	439	9	448
United Kingdom	228	9	237
France	161	10	171
Netherlands	139	21	160
Canada	95	4	99
Austria	76	9	85
Switzerland	75	3	78
Israel	56	3	59
Sweden	54	2	56
Australia	43	4	47
Italy	44	2	46
Belgium	42	1	43
Denmark	39	0	39
Japan	30	0	30
Spain	29	0	29
Poland	24	2	26
Finland	23	1	24
Czech Republic	16	0	16
Norway	13	3	16
Ireland	14	0	14
Brazil	11	0	11
Portugal	8	3	11
Singapore	11	0	11
Greece	10	0	10
India	10	0	10
Luxembourg	9	1	10
China	9	0	9
Hong Kong	6	0	6
South Africa	5	0	5
Hungary	4	0	4
New Zealand	3	1	4
Republic of Korea	4	0	4
Serbia	0	4	4
Slovenia	1	3	4
Chile	3	0	3
Cyprus	3	0	3
Philippines	3	0	3
Russian Federation	3	0	3
Tanzania	3	0	3
Argentina	2	0	2
Liechtenstein	2	0	2
Malta	2	0	2
Mexico	2	0	2
Romania	2	0	2
Turkey	2	0	2
Benin	1	0	1
Botswana	1	0	1
Croatia	1	0	1
Egypt	1	0	1
Estonia	0	1	1
Iceland	1	0	1
Qatar	1	0	1
Saudi Arabia	1	0	1
Taiwan	1	0	1
Uganda	1	0	1
United Arab Emirates	1	0	1
Total	2278	533	2811

(a) Details for 2022 by country



(b) Graphical distribution of seminars in group A in 2016–2022 by year and region



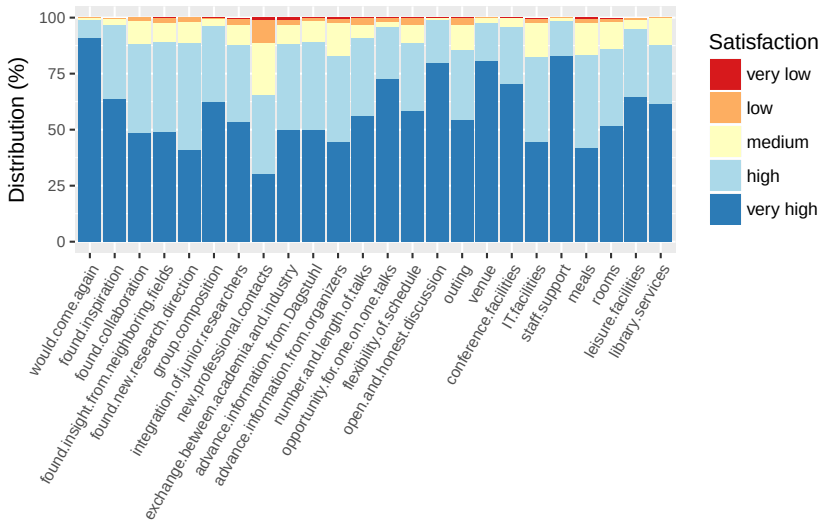
(c) Graphical distribution of events in group B in 2016–2022 by year and region

Region	Group A		Group B		Total	
	#	%	#	%	#	%
Europe (w/o Germany)	1025	45	75	14.1	1100	39.1
Germany	510	22.4	437	82	947	33.7
North America	534	23.4	13	2.4	547	19.5
Asia	133	5.8	3	0.6	136	4.8
Australia	46	2	5	0.9	51	1.8
South America	18	0.8	0	0	18	0.6
Africa	12	0.5	0	0	12	0.4
Total	2278	100	533	100	2811	100

(d) Details for 2022 by region

Fig. 13.11

Number of Dagstuhl guests by country of origin. Group A = Dagstuhl Seminar and Dagstuhl Perspectives Workshop participants. Group B = Participants in all other events (GI-Dagstuhl Seminars, educational events, and research group meetings).

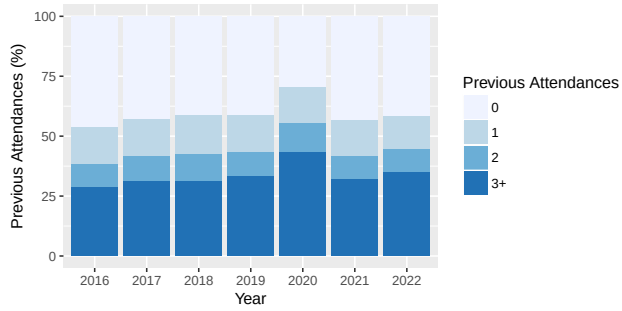


(a) Graphical distribution for 2022

	2016	2017	2018	2019	2020	2021	20222	2022 – Detailed Numbers					
	ø	ø	ø	ø	ø	ø	ø	1	2	3	4	5	total
would come again	4.9	4.9	4.9	4.9	4.9	4.9	4.9	0	3	8	90	1041	1142
found inspiration	4.5	4.5	4.6	4.6	4.6	4.5	4.6	0	5	32	375	730	1142
found collaboration	4.1	4.2	4.2	4.3	4.3	4.1	4.4	0	17	116	449	549	1131
found insight from neighboring fields	4.2	4.2	4.2	4.3	4.2	4.2	4.4	2	23	95	455	555	1130
found new research direction	4.1	4.1	4.2	4.2	4.2	4.2	4.3	0	20	107	532	461	1120
group composition	4.5	4.5	4.5	4.6	4.5	4.5	4.6	1	5	34	387	711	1138
integration of junior researchers	4.3	4.3	4.3	4.3	4.3	4.3	4.4	4	31	101	386	599	1121
new professional contacts	3.7	3.8	3.8	3.8	3.6	3.5	3.8	10	113	260	388	338	1109
exchange between academia and industry	4.3	4.4	4.4	4.4	4.3	4.3	4.3	7	15	62	273	359	716
advance information from Dagstuhl	4.4	4.4	4.4	4.5	4.5	4.3	4.4	1	14	106	433	551	1105
advance information from organizers	4.2	4.1	4.3	4.3	4.3	4.2	4.2	5	20	163	421	488	1097
number and length of talks	4.3	4.2	4.3	4.3	4.3	4.3	4.4	3	35	61	394	633	1126
opportunity for one on one talks	4.6	4.6	4.6	4.6	4.6	4.1	4.7	2	19	23	261	805	1110
flexibility of schedule	4.4	4.3	4.3	4.5	4.4	4.2	4.4	2	32	95	340	655	1124
open and honest discussion	4.7	4.7	4.7	4.7	4.7	4.6	4.8	1	1	10	218	907	1137
outing	4.2	4.2	4.3	4.2	4.2	4.4	4.4	2	24	97	272	470	865
venue	4.7	4.7	4.7	4.7	4.7	4.8	4.8	0	2	22	188	891	1103
conference facilities	4.7	4.7	4.7	4.7	4.7	4.6	4.7	1	1	42	279	770	1093
IT facilities	4.4	4.3	4.4	4.4	4.3	4.3	4.2	4	17	139	352	411	923
staff support	4.7	4.7	4.8	4.8	4.7	4.7	4.8	0	1	13	167	873	1054
meals	4.1	4.1	4.1	4.2	4.1	4.3	4.2	6	19	154	443	453	1075
rooms	4.4	4.4	4.4	4.4	4.3	4.4	4.4	4	15	133	368	555	1075
leisure facilities	4.5	4.5	4.6	4.6	4.6	4.5	4.6	0	8	43	299	649	999
library services	4.5	4.5	4.6	4.6	4.5	4.5	4.5	0	1	40	88	208	337

(b) Averages for 2016–2022 and detailed numbers for 2022: 1 = very low, 2 = low, 3 = medium, 4 = high, 5 = very high

Fig. 13.12
Satisfaction of Dagstuhl Seminar and Dagstuhl Perspectives Workshop participants, according to our guest survey.

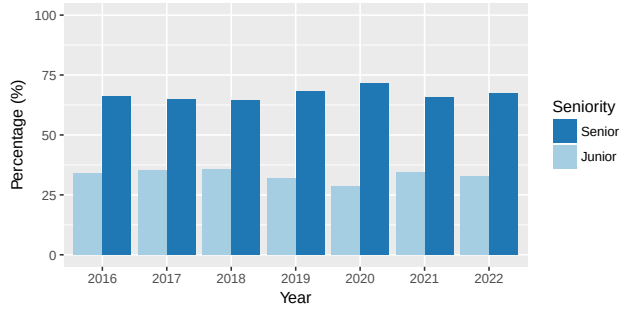


(a) Graphical distribution for 2016–2022

Year	Number of Previous Attendances								Total
	0		1		2		>2		
	#	%	#	%	#	%	#	%	#
2016	654	46	217	15	137	10	410	29	1418
2017	607	43	222	16	148	10	446	31	1423
2018	557	41	219	16	148	11	425	32	1349
2019	615	41	230	15	144	10	503	34	1492
2020	61	29	31	15	25	12	90	43	207
2021	283	43	96	15	63	10	211	32	653
2022	470	42	156	14	107	9	398	35	1131

(b) Detailed numbers for 2016–2022

Fig. 13.13
Dagstuhl Seminar and Dagstuhl Perspectives Workshop participants and their previous instances of attendance at Dagstuhl Seminars or Dagstuhl Perspectives Workshops, according to our guest survey.

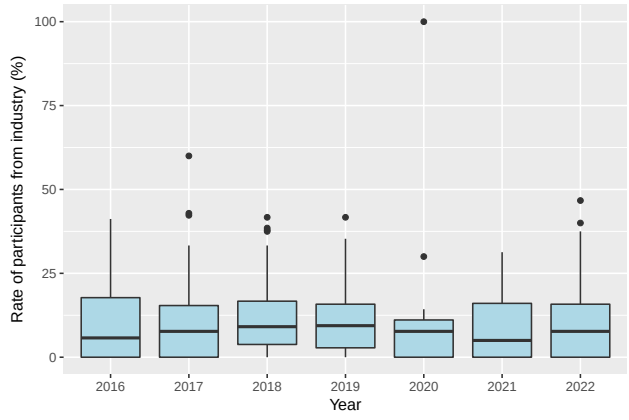


(a) Chart for 2016–2022

Year	Junior		Senior		Total
	#	%	#	%	#
2016	404	33.9	787	66.1	1191
2017	422	35.2	778	64.8	1200
2018	401	35.7	722	64.3	1123
2019	385	31.9	823	68.1	1208
2020	53	28.5	133	71.5	186
2021	195	34.5	370	65.5	565
2022	322	32.6	665	67.4	987

(b) Detailed numbers for 2016–2022

Fig. 13.14
Self-assigned seniority of Dagstuhl Seminar and Dagstuhl Perspectives Workshop participants, according to our guest survey.

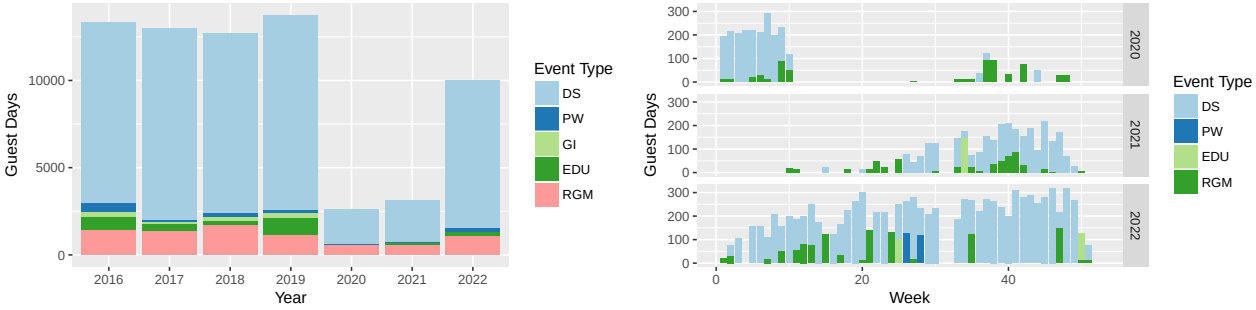


(a) Chart for 2016–2022

Year	Min (%)	Max (%)	Avg (%)	Std (%)
2016	0.0	41.2	10.3	11.0
2017	0.0	60.0	10.9	11.6
2018	0.0	41.7	11.1	10.4
2019	0.0	41.7	11.4	10.7
2020	0.0	100.0	14.6	25.9
2021	0.0	31.3	8.5	10.1
2022	0.0	46.7	10.1	11.1

(b) Detailed numbers for 2016–2022

Fig. 13.15
Distribution of the ratio of participants with self-assigned primary occupation in business per Dagstuhl Seminar and Dagstuhl Perspectives Workshop in 2016–2022, according to our guest survey. Min = minimal value, Max = maximal value, Avg = average, Std = standard deviation. Occupation in business includes "industrial research", "industrial development", and "self employed".



(a) Chart for 2016–2022

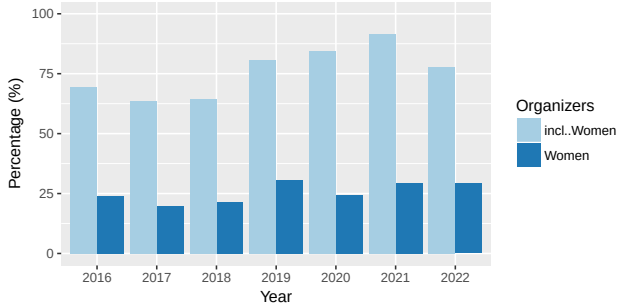
(b) Graphical distribution for 2020–2022 by week

Year	DS	PW	GI	EDU	RGM	Total
2016	10362	495	315	703	1462	13337
2017	10989	102	105	401	1391	12988
2018	10270	182	250	231	1740	12673
2019	11127	225	239	1004	1144	13739
2020	1984	0	0	0	614	2598
2021	2397	0	0	150	576	3123
2022	8435	252	0	218	1104	10009

(c) Detailed numbers for 2016–2022

Fig. 13.16

Number of overnight stays at Schloss Dagstuhl. DS = Dagstuhl Seminar, PW = Dagstuhl Perspectives Workshop, GI = GI-Dagstuhl Seminar, EDU = educational event, RGM = research group meeting.



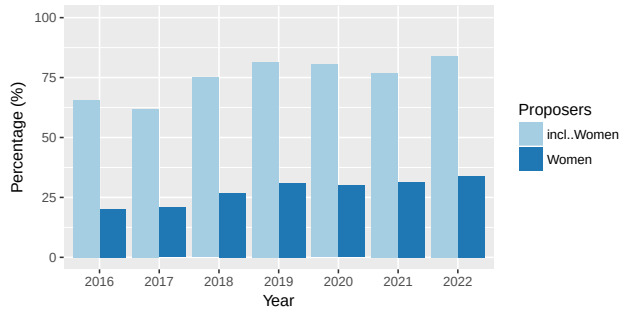
(a) Chart for 2016–2022

	Organizer Teams			Organizers		
Year	Total	incl. Women		Total	Women	
	#	#	%	#	#	%
2016	72	50	69.4	278	67	24.1
2017	77	49	63.6	303	60	19.8
2018	65	42	64.6	259	56	21.6
2019	73	59	80.8	291	89	30.6
2020	13	11	84.6	53	13	24.5
2021	47	43	91.5	184	54	29.3
2022	77	60	77.9	308	90	29.2

(b) Detailed numbers for 2016–2022

Fig. 13.17

Dagstuhl Seminars and Dagstuhl Perspectives Workshops with organizer teams including women.

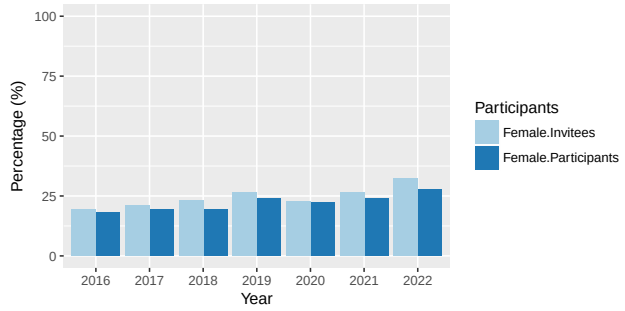


(a) Chart for 2016–2022

	Proposer Teams			Proposers		
Year	Total	incl. Women		Total	Women	
	#	#	%	#	#	%
2016	125	82	65.6	491	99	20.2
2017	102	63	61.8	394	82	20.8
2018	136	102	75.0	522	140	26.8
2019	103	84	81.6	411	127	30.9
2020	153	123	80.4	593	178	30.0
2021	74	57	77.0	296	93	31.4
2022	93	78	83.9	367	124	33.8

(b) Detailed numbers for 2016–2022

Fig. 13.18
Dagstuhl Seminar and Dagstuhl Perspectives Workshop proposals with proposer teams including women.

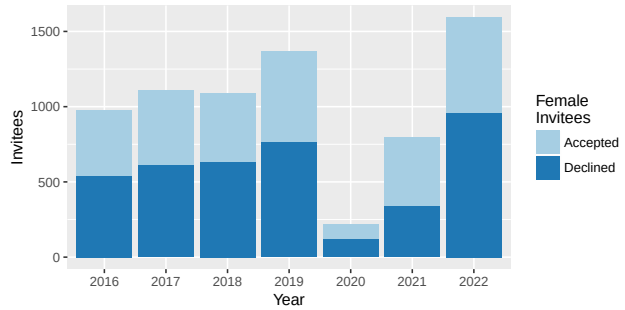


(a) Chart for 2016–2022

	Invitees			Participants		
Year	Total	Women		Total	Women	
	#	#	%	#	#	%
2016	5060	977	19.3	2393	437	18.3
2017	5267	1109	21.1	2572	494	19.2
2018	4692	1086	23.1	2320	453	19.5
2019	5143	1367	26.6	2498	603	24.1
2020	964	218	22.6	442	99	22.4
2021	3022	794	26.3	1894	456	24.1
2022	4950	1593	32.2	2278	635	27.9

(b) Detailed numbers for 2016–2022

Fig. 13.19
Female invitees and participants in Dagstuhl Seminars and Dagstuhl Perspectives Workshops, by year.

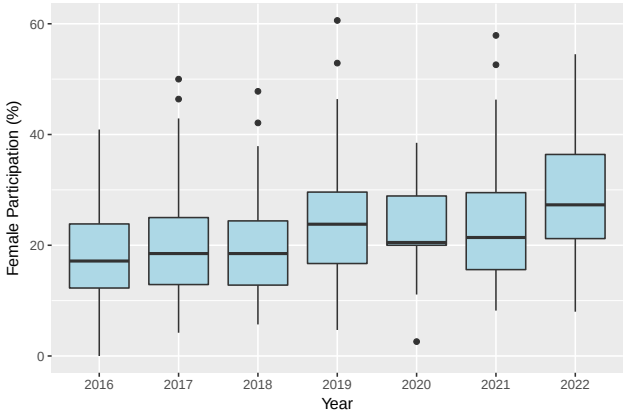


(a) Chart for 2016–2022

Year	Female Invitees	Accepted		Declined	
	#	#	%	#	%
2016	977	437	44.7	540	55.3
2017	1109	494	44.5	615	55.5
2018	1086	453	41.7	633	58.3
2019	1367	603	44.1	764	55.9
2020	218	99	45.4	119	54.6
2021	794	456	57.4	338	42.6
2022	1593	635	39.9	958	60.1

(b) Detailed numbers for 2016–2022

Fig. 13.20
Female invitees to Dagstuhl Seminars and Dagstuhl Perspectives Workshops.



(a) Chart for 2016–2022

Year	Min (%)	Max (%)	Avg (%)	Std (%)
2016	0.0	40.9	18.3	9.1
2017	4.2	50.0	19.6	9.8
2018	5.7	47.8	19.8	9.2
2019	4.7	60.6	24.6	10.5
2020	2.6	38.5	22.7	9.2
2021	8.2	57.9	24.5	11.4
2022	8.0	54.5	28.5	10.7

(b) Detailed numbers for 2016–2022

Fig. 13.21

Distribution of the share of women among participants per Dagstuhl Seminar or Dagstuhl Perspectives Workshop in 2016–2022. Min = minimal value, Max = maximal value, Avg = average, Std = standard deviation.

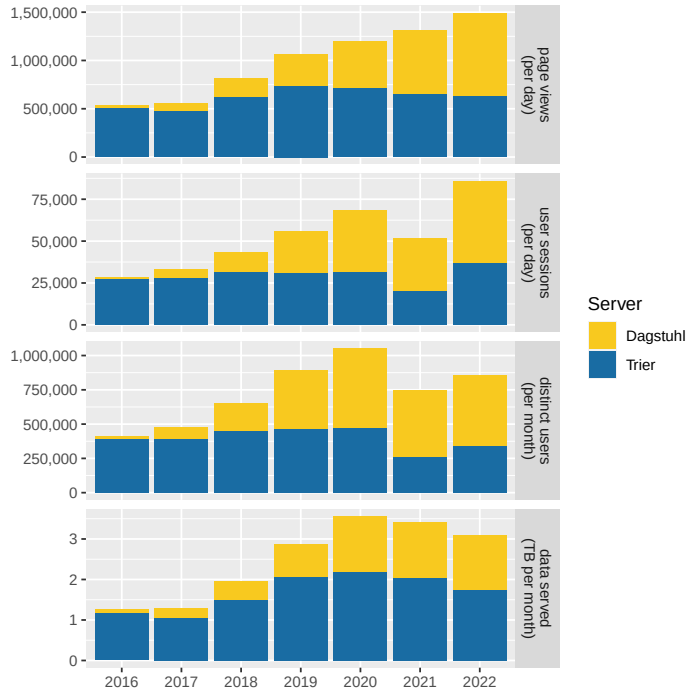
Statistiken zur Bibliographiedatenbank dblp

13.2

Statistics of the dblp computer science bibliography

Dieser Abschnitt enthält statistische Daten zur Bibliographiedatenbank dblp. Fig. 13.22 listet die durchschnittlichen Nutzungszahlen der letzten Jahre auf. Ein Überblick über die Entwicklung des dblp Datenbestandes kann Fig. 13.23 und Fig. 13.24 entnommen werden. Fig. 13.25–13.27 geben Auskunft über die kontinuierliche Datenkuration und -anreicherung des Bestandes.

This section provides statistical data about the dblp computer science bibliography. Fig. 13.22 shows the average usage statistics of the dblp servers in the past years. An overview of the development of the dblp database can be found in Fig. 13.23 and Fig. 13.24. Information about the continuous data curation and enrichment of existing records can be found in Fig. 13.25–13.27.

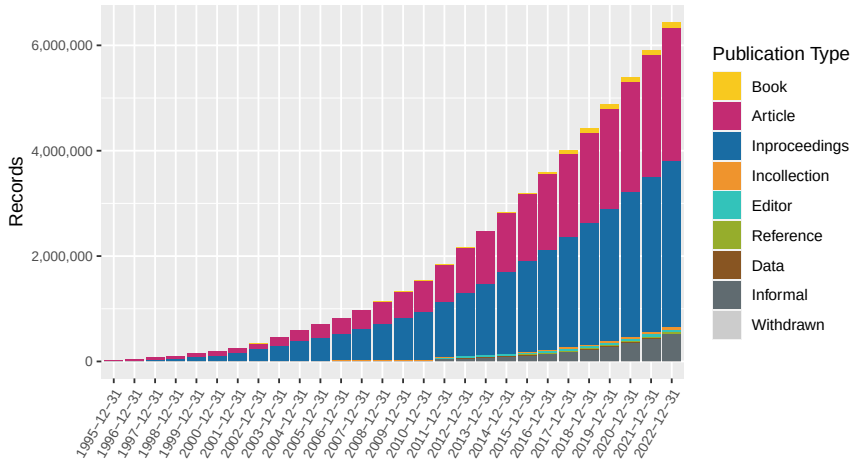


(a) Chart for 2016–2022

	Trier		Dagstuhl		Total		
	2021	2022	2021	2022	2021	2022	%
page views per day	644,837	627,039	672,274	866,072	1,317,111	1,493,111	+13.4
user sessions (visits) per day	20,134	36,639	31,537	49,453	51,672	86,093	+66.6
page views per user session	32.0	17.1	21.3	17.5	25.5	17.3	-32.0
distinct users (IPs) per month	257,769	335,436	492,268	517,785	750,038	853,222	+13.8
data served per month	2,073.9 GB	1,793.7 GB	1,433.5 GB	1,378.1 GB	3,507.4 GB	3,171.7 GB	-9.6

(b) Detailed numbers for the past two years

Fig. 13.22
Average usage of the two public dblp servers. Trier = dblp.uni-trier.de, Dagstuhl = dblp.dagstuhl.de. All figures exclude traffic caused by recognized bots and web crawlers. Since 2017, server Dagstuhl has been promoted to play a more prominent role under the domain dblp.org. In 2021, the counting method to determine distinct users has been changed in order to avoid double counting of the same IP across diferent servers.

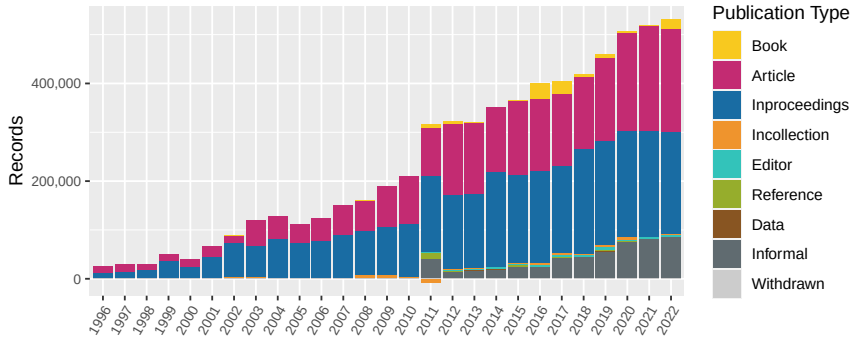


(a) Chart for 1996–2022

Year	Book		Article		Inproceedings		Incollection		Editor		Reference		Data		Informal		Withdrawn		Total
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#
2016	51,070	1.4	1,429,427	39.7	1,912,895	53.1	19,774	0.5	33,782	0.9	20,174	0.6	26	0.0	134,354	3.7	0	0.0	3,601,502
2017	77,408	1.9	1,576,972	39.4	2,091,486	52.2	23,101	0.6	37,049	0.9	23,089	0.6	49	0.0	174,723	4.4	1,609	0.0	4,005,486
2018	83,249	1.9	1,725,704	39.0	2,306,585	52.1	24,708	0.6	40,795	0.9	23,150	0.5	514	0.0	216,984	4.9	3,036	0.1	4,424,725
2019	90,826	1.9	1,896,014	38.8	2,518,298	51.6	30,457	0.6	44,898	0.9	26,997	0.6	1,402	0.0	271,633	5.6	3,494	0.1	4,884,019
2020	95,442	1.8	2,096,005	38.9	2,735,716	50.7	36,372	0.7	48,481	0.9	27,321	0.5	2,230	0.0	345,676	6.4	3,827	0.1	5,391,070
2021	99,289	1.7	2,310,800	39.1	2,951,745	49.9	37,745	0.6	51,968	0.9	27,321	0.5	2,882	0.0	425,791	7.2	4,126	0.1	5,911,667
2022	119,483	1.9	2,522,724	39.2	3,160,762	49.1	39,407	0.6	55,695	0.9	27,321	0.4	3,543	0.1	508,577	7.9	5,847	0.1	6,443,359

(b) Detailed numbers for 2016–2022

Fig. 13.23
Development of the total size of the dblp database.

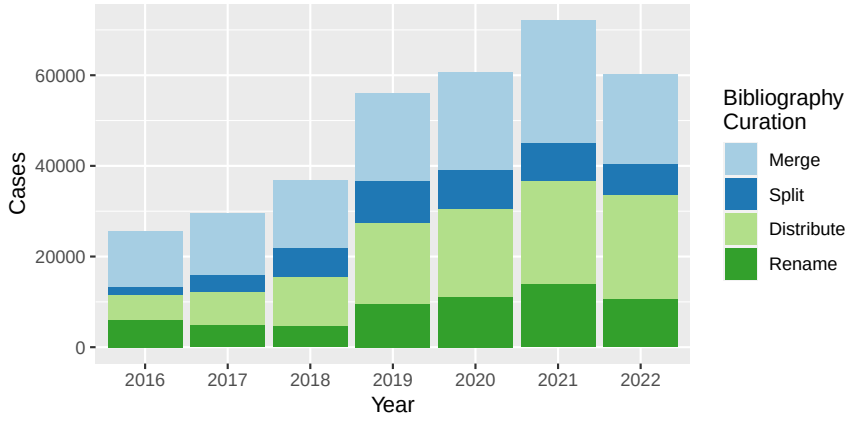


(a) Chart for 1996–2022

Year	Book		Article		Inproceedings		Incollection		Editor		Reference		Data		Informal		Withdrawn		Total
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#
2016	32,752	8.2	148,182	36.9	188,633	47.0	3,486	0.9	3,738	0.9	1,071	0.3	14	0.0	23,380	5.8	0	0.0	401,256
2017	26,338	6.5	147,545	36.5	178,591	44.2	3,327	0.8	3,267	0.8	2,915	0.7	23	0.0	40,369	10.0	1,609	0.4	403,984
2018	5,841	1.4	148,732	35.5	215,099	51.3	1,607	0.4	3,746	0.9	61	0.0	465	0.1	42,261	10.1	1,427	0.3	419,239
2019	7,577	1.6	170,310	37.1	211,713	46.1	5,749	1.3	4,103	0.9	3,847	0.8	888	0.2	54,649	11.9	458	0.1	459,294
2020	4,616	0.9	199,991	39.4	217,418	42.9	5,915	1.2	3,583	0.7	324	0.1	828	0.2	74,043	14.6	333	0.1	507,051
2021	3,847	0.7	214,795	41.3	216,029	41.5	1,373	0.3	3,487	0.7	0	0.0	652	0.1	80,115	15.4	299	0.1	520,597
2022	20,194	3.8	211,924	39.9	209,017	39.3	1,662	0.3	3,727	0.7	0	0.0	661	0.1	82,786	15.6	1,721	0.3	531,692

(b) Detailed numbers for 2016–2022

Fig. 13.24
Development of newly included publications in dblp. The negative number of new *Incollection* records in 2011 results from relabeling several thousand existing records with the newly introduced *Reference* type. Similarly, in the same year, several thousand *Articles* and *Inproceedings* records have been labeled as *Informal*.



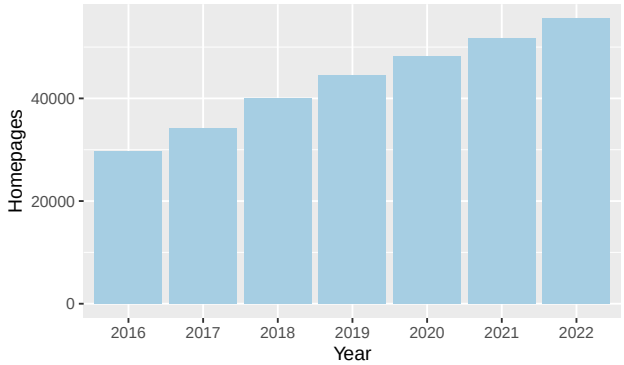
(a) Chart for 2016–2022

Year	Merge		Split		Distribute		Rename		Total
	#	%	#	%	#	%	#	%	#
2016	12,426	48.4	1,913	7.5	5,310	20.7	5,999	23.4	25,648
2017	13,537	46.0	3,660	12.4	7,465	25.3	4,786	16.3	29,448
2018	14,906	40.6	6,282	17.1	11,014	30.0	4,524	12.3	36,726
2019	19,595	34.9	9,192	16.4	17,795	31.7	9,562	17.0	56,144
2020	21,636	35.7	8,606	14.2	19,326	31.9	11,083	18.3	60,651
2021	27,204	37.7	8,254	11.4	22,877	31.7	13,785	19.1	72,120
2022	19,739	32.8	6,813	11.3	23,065	38.4	10,508	17.5	60,125

(b) Detailed numbers for 2016–2022

Fig. 13.25

Curation of existing dblp author bibliographies. The figures give the number of distinct edit cases (measured between the first and the last day of every given year) where a dblp team member manually corrected the assignment of publications within dblp author bibliographies. We distinguish between four curation cases: *Merge* = Two or more synonymous bibliographies have been merged into a single bibliography. *Split* = A single, homonymous bibliography has been split into two or more bibliographies. *Distribute* = A mixed case where records from two or more bibliographies have been redistributed between two or more bibliographies. *Rename* = A case where no actual publications have been reassigned, but the surface form of the author name(s) of a bibliography have been corrected or improved.



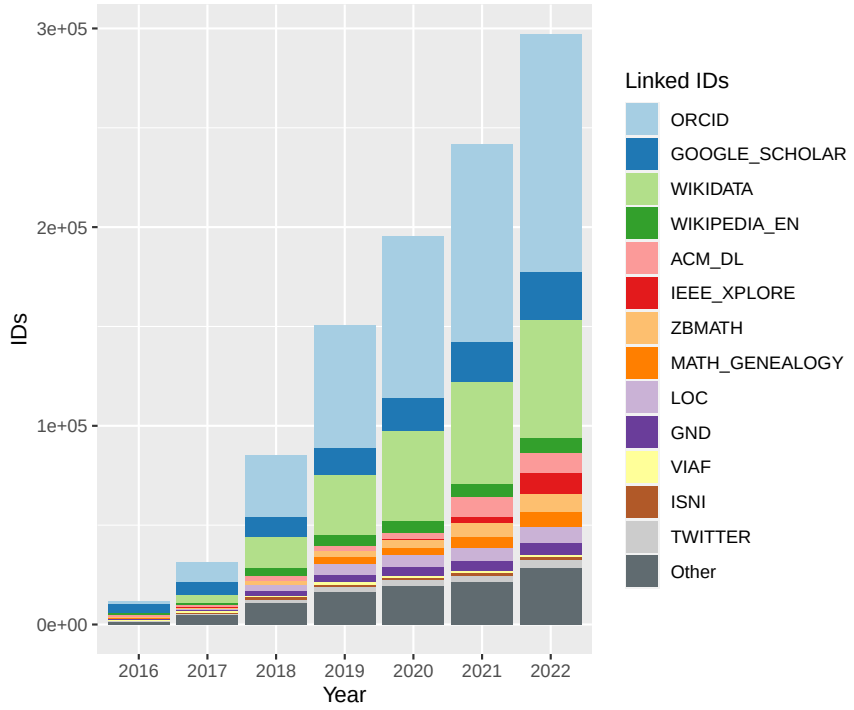
(a) Chart for 2016–2022

Year	Homepages
2016	29,640
2017	34,281
2018	40,018
2019	44,519
2020	48,216
2021	51,756
2022	55,572

(b) Detailed numbers for 2016–2022

Fig. 13.26

Linked and verified academic homepages in dblp author bibliographies. A single author bibliography may be linked to multiple academic homepages. These figures exclude linked external IDs which are given in Figure 13.27.



(a) Chart for 2016–2022

Year	ORCID		Google Scholar		Wikidata		Wikipedia (en)		ACM DL		IEEE Xplore		zbMATH	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%
2016	1,420	12.2	4,522	38.9	4	0.0	1,103	9.5	1,107	9.5	0	0.0	633	5.5
2017	10,342	32.8	6,608	21.0	3,588	11.4	1,376	4.4	1,277	4.1	0	0.0	646	2.0
2018	30,939	36.4	10,182	12.0	15,565	18.3	4,366	5.1	2,144	2.5	0	0.0	2,145	2.5
2019	61,976	41.1	13,726	9.1	30,022	19.9	5,547	3.7	2,448	1.6	104	0.1	3,363	2.2
2020	81,675	41.8	16,338	8.4	45,216	23.1	6,275	3.2	3,166	1.6	557	0.3	3,607	1.8
2021	99,459	41.2	20,060	8.3	51,744	21.4	6,668	2.8	9,642	4.0	3,013	1.2	7,486	3.1
2022	119,683	40.3	24,443	8.2	59,514	20.0	7,312	2.5	10,306	3.5	10,269	3.5	9,524	3.2

Year	Math Genealogy		LOC		GND		VIAF		ISNI		Twitter		Other		Total
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#
2016	84	0.7	324	2.8	220	1.9	526	4.5	7	0.1	161	1.4	1,503	12.9	11,614
2017	5	0.0	602	1.9	430	1.4	840	2.7	606	1.9	392	1.2	4,804	15.2	31,516
2018	65	0.1	2,862	3.4	2,385	2.8	938	1.1	1,231	1.4	1,567	1.8	10,670	12.5	85,059
2019	3,071	2.0	5,680	3.8	3,782	2.5	1,060	0.7	1,383	0.9	2,183	1.4	16,349	10.8	150,694
2020	3,590	1.8	6,289	3.2	4,241	2.2	1,061	0.5	1,385	0.7	2,613	1.3	19,358	9.9	195,371
2021	5,048	2.1	6,877	2.8	4,777	2.0	1,068	0.4	1,390	0.6	3,183	1.3	21,233	8.8	241,648
2022	7,060	2.4	8,234	2.8	5,938	2.0	1,069	0.4	1,399	0.5	3,786	1.3	28,652	9.6	297,189

(b) Detailed numbers for 2016–2022

Fig. 13.27
Linked and verified external person IDs in dblp author bibliographies. A single bibliography may be linked to multiple external IDs.

Statistiken zu Dagstuhl Publishing

Dieser Abschnitt enthält statistische Daten zum Publikationswesen von Schloss Dagstuhl.

Ein Überblick über die Entwicklung der seminarbezogenen Veröffentlichungen kann den ersten drei Diagrammen und Tabellen entnommen werden. Fig. 13.28 fasst die statistischen Daten der Veröffentlichungen in der Zeitschrift Dagstuhl Reports zusammen, Fig. 13.29 die der Publikationen in der Reihe Dagstuhl Manifestos und schließlich Fig. 13.30 die der veröffentlichten Bände in der Reihe Dagstuhl Follow-Ups.

Die statistischen Daten zu den dienstleistungsbezogenen Veröffentlichungen finden sich anschließend: Fig. 13.31 fasst die Daten in der Reihe OASIcs und Fig. 13.32 die der Reihe LIPIcs zusammen.

Die Kennzahlen der Zeitschrift LITES können Fig. 13.33 entnommen werden.

Die verschiedenen Publikationsserien wurden in unterschiedlichen Jahren zwischen 2009 und 2015 gegründet. Wir stellen in den Statistiken dennoch stets den gesamten Zeitraum (2016–2022) dar.

13.3 Statistics of Dagstuhl Publishing

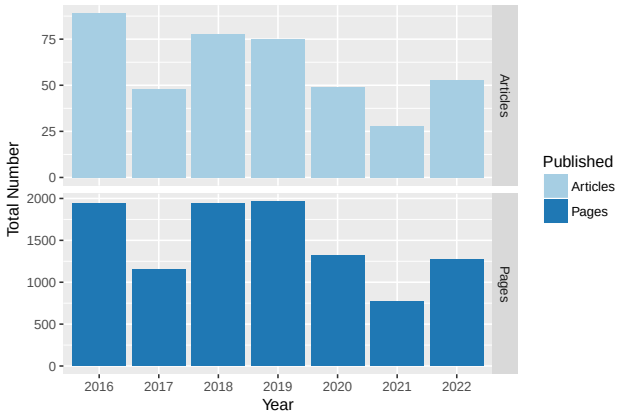
In this section the statistical data of Dagstuhl Publishing is presented.

The first three figures present the development of the seminar-focused series: Fig. 13.28 summarizes the data of the periodical Dagstuhl Reports, Fig. 13.29 the data of the Dagstuhl Manifestos series, and, finally, Fig. 13.30 that of the volumes published in the Dagstuhl Follow-Ups series.

The statistical data of the service-focused series are presented afterwards. Fig. 13.31 presents numbers related to OASIcs and Fig. 13.32 numbers related to LIPIcs.

We summarize the publications of the journal LITES in Fig. 13.33.

Please note that the publication series were established in different years in the period between 2009 and 2015. However, we always consider this complete period (2016–2022).

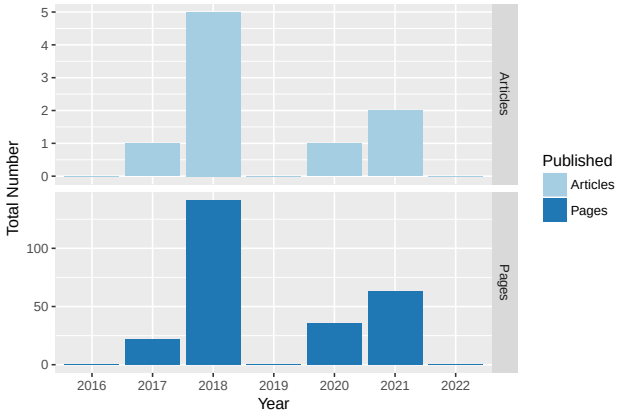


(a) Graphical distribution for 2016–2022

Year	Articles	Pages
2016	89	1940
2017	48	1154
2018	78	1938
2019	75	1959
2020	49	1322
2021	28	770
2022	53	1276

(b) Detailed numbers for 2016–2022

Fig. 13.28
Statistics about Dagstuhl Reports published between 2016 to 2022.

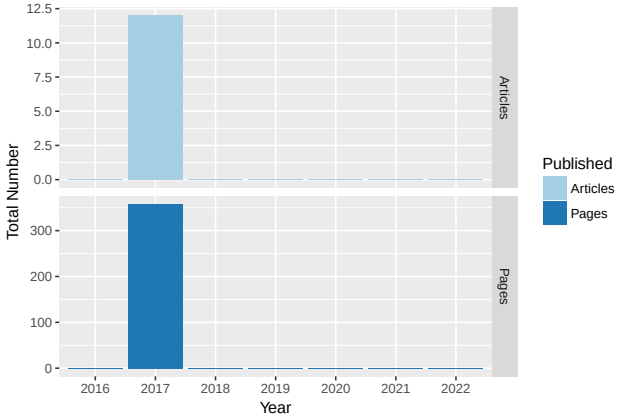


(a) Graphical distribution for 2016–2022

Year	Articles	Pages
2016	0	0
2017	1	22
2018	5	141
2019	0	0
2020	1	36
2021	2	63
2022	0	0

(b) Detailed numbers for 2016–2022

Fig. 13.29
Statistics about Dagstuhl Manifestos published between 2016 to 2022.

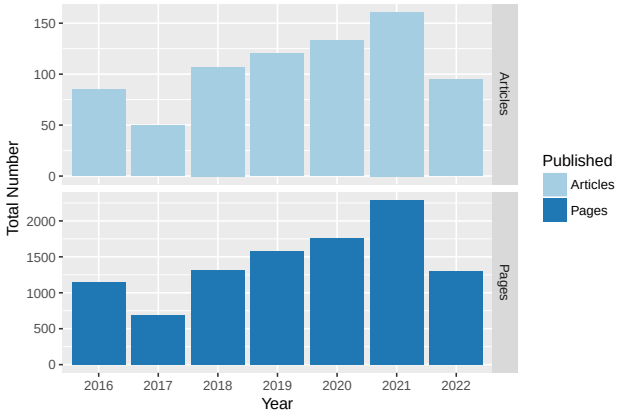


(a) Graphical distribution for 2016–2022

Year	Volumes	Articles	Pages
2016	0	0	0
2017	1	12	358
2018	0	0	0
2019	0	0	0
2020	0	0	0
2021	0	0	0
2022	0	0	0

(b) Detailed numbers for 2016–2022

Fig. 13.30
Statistics about Dagstuhl Follow-Ups volumes published between 2016 to 2022.

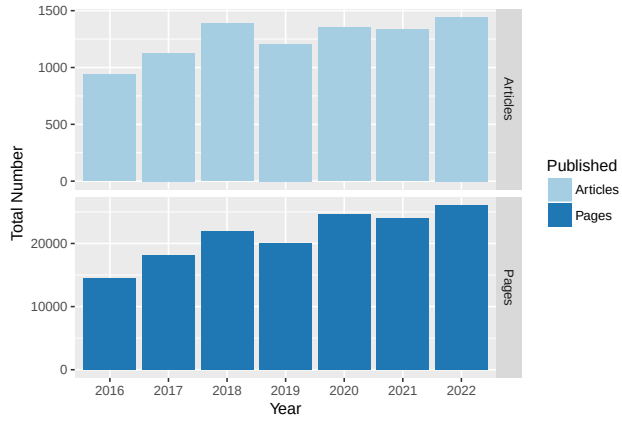


(a) Graphical distribution for 2016–2022

Year	Volumes	Articles	Pages
2016	6	85	1152
2017	3	50	694
2018	7	107	1312
2019	9	120	1576
2020	11	133	1754
2021	11	161	2284
2022	10	95	1306

(b) Detailed numbers for 2016–2022

Fig. 13.31
Statistics about OASIcs volumes published between 2016 to 2022.

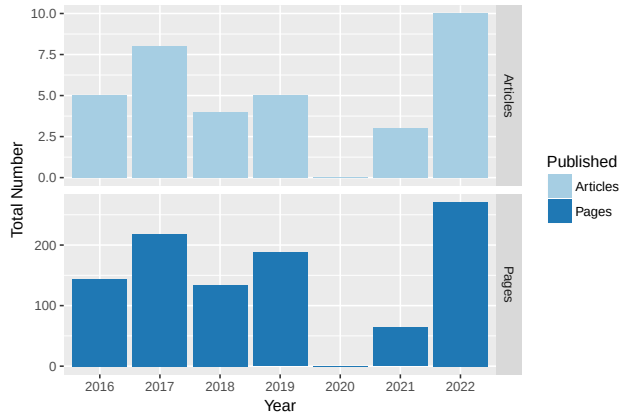


(a) Graphical distribution for 2016–2022

Year	Volumes	Articles	Pages
2016	19	939	14538
2017	25	1127	18082
2018	32	1387	21876
2019	29	1208	20032
2020	32	1352	24562
2021	32	1333	23960
2022	36	1444	25998

(b) Detailed numbers for 2016–2022

Fig. 13.32
Statistics about LIPIcs volumes published between 2016 to 2022.

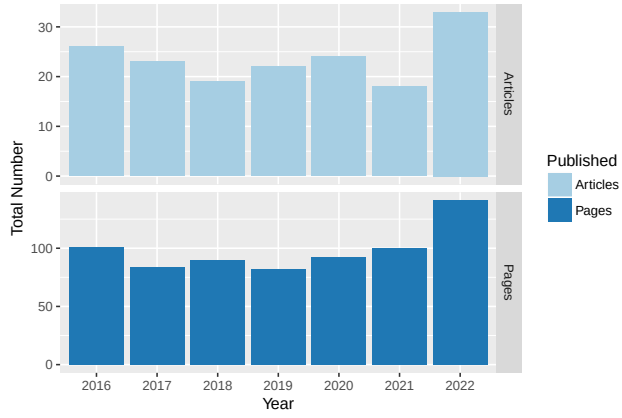


(a) Graphical distribution for 2016–2022

Year	Articles	Pages
2016	5	144
2017	8	218
2018	4	134
2019	5	188
2020	0	0
2021	3	64
2022	10	271

(b) Detailed numbers for 2016–2022

Fig. 13.33
Statistics about LITES articles published between 2016 to 2022.



(a) Graphical distribution for 2016–2022

Year	Articles	Pages
2016	26	101
2017	23	84
2018	19	90
2019	22	82
2020	24	92
2021	18	100
2022	33	141

(b) Detailed numbers for 2016–2022

Fig. 13.34
Statistics about DARTS artifacts published between 2016 to 2022.

14 Veranstaltungen 2022
Schedule of Events 2022

Dagstuhl-Seminare 14.1 *Dagstuhl Seminars*

22021 – Mobility Data Science
Mohamed Mokbel (University of Minnesota – Minneapolis, US), Mahmoud Sakr (UL – Brussels, BE), Li Xiong (Emory University – Atlanta, US), Andreas Züfle (George Mason Univ. – Fairfax, US)
January 9–14, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22021

22022 – Mobility Data Analysis: From Technical to Ethical
Bettina Berendt (TU Berlin, DE), Stan Matwin (Dalhousie University – Halifax, CA), Chiara Renso (ISTI-CNR – Pisa, IT)
January 9–12, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22022

22031 – Bringing Graph Databases and Network Visualization Together
Karsten Klein (Universität Konstanz, DE), Juan F. Sequeda (data.world – Austin, US), Hsiang-Yun Wu (FH – St. Pölten , AT), Da Yan (The University of Alabama – Birmingham, US)
January 16–21, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22031

22042 – Privacy Protection of Automated and Self-Driving Vehicles
Frank Kargl (Universität Ulm, DE), Ioannis Krontiris (Huawei Technologies – München, DE), André Weimerskirch (Lear Corporation – Ann Arbor, US), Ian Williams (University of Michigan – Ann Arbor, US)
January 23–28, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22042

22051 – Finite and Algorithmic Model Theory
Albert Atserias (UPC Barcelona Tech, ES), Christoph Berkholz (HU Berlin, DE), Kousha Etessami (University of Edinburgh, GB), Joanna Ochremiak (University of Bordeaux, FR)
January 30 to February 4, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22051

22052 – The Human Factors Impact of Programming Error Messages
Brett A. Becker (University College Dublin, IE), Paul Denny (University of Auckland, NZ), Janet Siegmund (TU Chemnitz, DE), Andreas Stefik (University of Nevada – Las Vegas, US)
January 30 to February 2, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22052

22061 – Logic and Random Discrete Structures
Erich Grädel (RWTH Aachen, DE), Phokion G. Kolaitis (University of California – Santa Cruz, US), Marc Noy (UPC Barcelona Tech, ES)
February 6–11, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22061

22062 – Computation and Reconfiguration in Low-Dimensional Topological Spaces
Maike Buchin (Ruhr-Universität Bochum, DE), Arnaud de Mesmay (University Paris-Est – Marne-la-Vallée, FR), Anna Lubiw (University of Waterloo, CA), Saul Schleimer (University of Warwick – Coventry, GB)
February 6–11, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22062

22072 – New Perspectives in Symbolic Computation and Satisfiability Checking
Erika Abraham (RWTH Aachen University, DE), James H. Davenport (University of Bath, GB), Matthew England (Coventry University, GB), Alberto Griggio (Bruno Kessler Foundation – Trento, IT)
February 13–18, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22072

22081 – Theory of Randomized Optimization Heuristics
Anne Auger (INRIA Saclay – Palaiseau, FR), Carlos M. Fonseca (University of Coimbra, PT), Tobias Friedrich (Hasso-Plattner-Institut, Universität Potsdam, DE), Johannes Lengler (ETH Zürich, CH)
February 20–25, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22081

22082 – Deep Learning and Knowledge Integration for Music Audio Analysis
Rachel Bittner (Spotify – Paris, FR), Meinard Müller (Universität Erlangen-Nürnberg, DE), Juhan Nam (KAIST – Daejeon, KR)
February 20–25, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22082

22091 – AI for the Social Good
Claudia Clopath (Imperial College London, GB), Ruben De Winne (Oxfam Novib – The Hague, NL), Rayid Ghani (Carnegie Mellon University – Pittsburgh, US), Tom Schaul (Google DeepMind – London, GB)
February 27 to March 4, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22091

22101 – Tensor Computations: Applications and Optimization
Paolo Bientinesi (University of Umeå, SE), David Ham (Imperial College London, GB), Furong Huang (University of Maryland – College Park, US), Paul H. J. Kelly (Imperial College London, GB), P. (Saday) Sadayappan (University of Utah – Salt Lake City, US)
March 6–11, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22101

22102 – Computational Models of Human-Automated Vehicle Interaction
Martin Baumann (Universität Ulm, DE), Shamsi Tamara Iqbal (Microsoft – Redmond, US), Christian P. Janssen (Utrecht University, NL), Antti Oulasvirta (Aalto University, FI)
March 6–11, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22102

22111 – Database Indexing and Query Processing
Renata Borovica-Gajic (The University of Melbourne, AU), Goetz Graefe (Google – Madison, US), Allison Lee (Snowflake – San Mateo, US), Caetano Sauer (Salesforce – München, DE), Pinar Tözün (IT University of Copenhagen, DK)
March 13–18, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22111

22121 – 3D Morphable Models and Beyond
Bernhard Egger (Universität Erlangen-Nürnberg, DE), William Smith (University of York, GB), Christian Theobalt (MPI für Informatik – Saarbrücken, DE), Stefanie Wuhrer (INRIA – Grenoble, FR)
March 20–25, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22121

22131 – Framing in Communication: From Theories to Computation
Katarzyna Budzynska (Warsaw University of Technology, PL), Chris Reed (University of Dundee, GB), Manfred Stede (Universität Potsdam, DE), Benno Stein (Bauhaus-Universität Weimar, DE)
March 27 to April 1, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22131

22132 – Graph Embeddings: Theory Meets Practice
Martin Grohe (RWTH Aachen University, DE), Stephan Günnemann (TU München, DE), Stefanie Jegelka (MIT – Cambridge, US), Christopher Morris (McGill University & MILA – Montreal)
March 27–30, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22132

22141 – Symmetric Cryptography
Nils Gregor Leander (Ruhr-Universität Bochum, DE), Bart Mennink (Radboud University Nijmegen, NL), Maria Naya-Plasencia (INRIA – Paris, FR), Yu Sasaki (NTT – Tokyo, JP)
April 3–8, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22141

22161 – Recent Advancements in Tractable Probabilistic Inference
Priyank Jaini (Google – Toronto, CA), Kristian Kersting (TU Darmstadt, DE), Antonio Vergari (University of Edinburgh, GB), Max Welling (University of Amsterdam, NL)
April 18–22, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22161

22162 – Urban Mobility Analytics
David Jonietz (HERE – Zürich, CH), Monika Sester (Leibniz Universität Hannover, DE), Kathleen Stewart (University of Maryland – College Park, US), Martin Tomko (University of Melbourne – Carlton, AU), Stephan Winter (The University of Melbourne, AU)
April 18–22, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22162

22171 – Digital Twins for Cyber-Physical Systems Security
Mohammad Al Faruque (University of California – Irvine, US), Alvaro Cárdenas Mora (University of California – Santa Cruz, US), Simin Nadjm-Tehrani (Linköping University, SE), Edgar Weippl (Universität Wien, AT)
April 24–29, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22171

22172 – Technologies to Support Critical Thinking in an Age of Misinformation
Andreas Dengel (DFKI – Kaiserslautern, DE), Laurence Devillers (CNRS – Orsay, FR & Sorbonne University – Paris, FR), Tilman Dingler (The University of Melbourne, AU), Koichi Kise (Osaka Prefecture University, JP), Benjamin Tag (The University of Melbourne, AU)
April 24–27, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22172

22181 – Computational Metabolomics: From Spectra to Knowledge
Corey Broeckling (Colorado State University – Fort Collins, US), Timothy Ebbels (Imperial College London, GB), Ewy Mathé (National Institutes of Health – Bethesda, US), Nicola Zamboni (ETH Zürich, CH)
May 1–6, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22181

22182 – Estimation-of-Distribution Algorithms: Theory and Applications
Josu Ceberio Uribe (University of the Basque Country – Donostia, ES), Benjamin Doerr (Ecole Polytechnique – Palaiseau, FR), Fernando Lobo (University of Algarve, PT), Carsten Witt (Technical University of Denmark – Lyngby, DK)
May 1–6, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22182

22191 – Visual Text Analytics
Christopher Collins (Ontario Tech – Oshawa, CA), Antske Fokkens (Free University Amsterdam, NL), Andreas Kerren (Linköping University, SE), Chris Weaver (University of Oklahoma – Norman, US)
May 8–13, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22191

22192 – Dynamic Traffic Models in Transportation Science
Martin Gairing (University of Liverpool, GB), Carolina Osorio (HEC Montréal, CA & Google – Mountain View, US), Britta Peis (RWTH Aachen, DE), David Watling (University of Leeds, GB)
May 8–13, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22192

22201 – The Constraint Satisfaction Problem: Complexity and Approximability
Martin Grohe (RWTH Aachen University, DE), Venkatesan Guruswami (Carnegie Mellon University – Pittsburgh, US), Dániel Marx (CISPA – Saarbrücken, DE), Stanislav Živný (University of Oxford, GB)
May 15–20, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22201

22202 – Anticipatory Human-Machine Interaction
Jelmer Borst (University of Groningen, NL), Andreas Bulling (Universität Stuttgart, DE), Cleotilde Gonzalez (Carnegie Mellon University – Pittsburgh, US), Nele Rußwinkel (TU Berlin, DE)
May 15–20, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22202

22221 – Exponential Analysis: Theoretical Progress and Technological Innovation
Annie Cuyt (University of Antwerp, BE), Wen-shin Lee (University of Stirling, GB), Gerlind Plonka-Hoch (Universität Göttingen, DE)
May 29 to June 3, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22221

22222 – Radical Innovation and Design for Connected and Automated Vehicles
Wendy Ju (Cornell Tech – New York, US), Bastian Pfleging (TU Bergakademie Freiberg, DE), Andreas Riener (TH Ingolstadt, DE)
May 29 to June 3, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22222

22231 – Theories of Programming
Amy Ko (University of Washington – Seattle, US), Thomas D. LaToza (George Mason University – Fairfax, US), David C. Shepherd (Virginia Commonwealth University – Richmond, US), Dag Sjøberg (University of Oslo, NO)
June 6–10, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22231

22232 – Efficient and Equitable Natural Language Processing in the Age of Deep Learning
Jesse Dodge (AI2 – Seattle, US), Iryna Gurevych (TU Darmstadt, DE), Roy Schwartz (The Hebrew University of Jerusalem, IL), Emma Strubell (Carnegie Mellon University – Pittsburgh, US)
June 6–10, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22232

22251 – Human-Game AI Interaction
Dan Ashlock (in memoriam; † April 5, 2022), Setareh Maghsudi (Universität Tübingen, DE), Diego Perez Liebana (Queen Mary University of London, GB), Pieter Spronck (Tilburg University, NL)
June 19–24, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22251

22261 – Visualization Empowerment: How to Teach and Learn Data Visualization
Benjamin Bach (University of Edinburgh, GB), Sheelagh Carpendale (Simon Fraser University – Burnaby, CA), Uta Hinrichs (University of Edinburgh, GB), Samuel Huron (Institut Polytechnique de Paris, FR)
June 26 to July 1, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22261

22271 – Algorithms for Participatory Democracy
Markus Brill (TU Berlin, DE), Jiehua Chen (TU Wien, AT), Andreas Darmann (Universität Graz, AT), David Pennock (Rutgers University – Piscataway, US)
July 3–8, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22271

22272 – Eat-IT: Interactive Food
Masahiko Inami (University of Tokyo, JP), Sohyeong Kim (Stanford University, US), Florian 'Floyd' Mueller (Monash University – Clayton, AU), Marianna Obrist (University College London, GB)
July 3–8, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22272

22281 – Security of Machine Learning
Battista Biggio (University of Cagliari, IT), Nicholas Carlini (Google – Mountain View, US), Pavel Laskov (Universität Liechtenstein, LI), Konrad Rieck (TU Braunschweig, DE)
July 10–15, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22281

22291 – Machine Learning and Logical Reasoning: The New Frontier
Sébastien Bardin (CEA LIST, FR), Vijay Ganesh (University of Waterloo, CA), Somesh Jha (University of Wisconsin-Madison, US), Joao Marques-Silva (CNRS – Toulouse, FR)
July 17–22, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22291

22292 – Computational Approaches for Digitized Historical Newspapers
Antoine Doucet (University of La Rochelle, FR), Marten Düring (University of Luxembourg, LU), Maud Ehrmann (EPFL – Lausanne, CH), Clemens Neudecker (Staatsbibliothek zu Berlin, DE)
July 17–22, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22292

22301 – Algorithmic Aspects of Information Theory
Phokion G. Kolaitis (University of California – Santa Cruz, US & IBM Research, US), Andrei Romashchenko (University of Montpellier – LIRMM, FR & CNRS, FR), Milan Studený (The Czech Academy of Sciences – Prague, CZ), Dan Suciu (University of Washington – Seattle, US)
July 24–29, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22301

22302 – Educational Programming Languages and Systems
Neil Brown (King's College London, GB), Mark J. Guzdial (University of Michigan – Ann Arbor, US), Shriram Krishnamurthi (Brown University – Providence, US), Jens Mönig (SAP SE – Walldorf, DE)
July 24–29, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22302

22331 – Visualization and Decision Making Design Under Uncertainty
Nadia Boukhelifa (INRAE – Palaiseau, FR), Christopher R. Johnson (University of Utah – Salt Lake City, US), Kristi Potter (NREL – Golden, US)
August 15–19, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22331

22332 – Differential Equations and Continuous-Time Deep Learning
David Duvenaud (University of Toronto, CA), Markus Heinonen (Aalto University, FI), Michael Tiemann (Robert Bosch GmbH – Renningen, DE), Max Welling (University of Amsterdam, NL)
August 15–19, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22332

22341 – Power and Energy-Aware Computing on Heterogeneous Systems (PEACHES)
Julian De Hoog (The University of Melbourne, AU), Kerstin I. Eder (University of Bristol, GB), Timo Hönig (Ruhr-Universität Bochum, DE), Daniel Mosse (University of Pittsburgh, US), Max Plauth (Hasso-Plattner-Institut, Universität Potsdam, DE)
August 21–26, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22341

22342 – Privacy in Speech and Language Technology
Simone Fischer-Hübner (Karlstad University, SE), Dietrich Klakow (Universität des Saarlandes, DE), Peggy Valcke (KU Leuven, BE), Emmanuel Vincent (INRIA Nancy – Grand Est, FR)
August 21–26, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22342

22351 – Interactive Visualization for Fostering Trust in ML
Polo Chau (Georgia Institute of Technology – Atlanta, US), Alex Endert (Georgia Institute of Technology – Atlanta, US), Daniel A. Keim (Universität Konstanz, DE), Daniela Oelke (Hochschule Offenburg, DE)
August 28 to September 2, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22351

22361 – Challenges and Opportunities of Democracy in the Digital Society
Abraham Bernstein (Universität Zürich, CH), Anita Gohdes (Hertie School of Governance – Berlin, DE), Beth Simone Noveck (New York University, US), Steffen Staab (Universität Stuttgart, DE)
September 4–9, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22361

22362 – Model-Driven Engineering of Digital Twins
Benoit Combemale (University & IRISA – Rennes, FR), Vinay Kulkarni (Tata Consultancy Services – Pune, IN), Bernhard Rumpe (RWTH Aachen, DE), Steffen Zschaler (King's College London, GB)
September 4–9, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22362

22371 – Algebraic and Analytic Methods in Computational Complexity
Markus Bläser (Universität des Saarlandes – Saarbrücken, DE), Valentine Kabanets (Simon Fraser University – Burnaby, CA), Ronen Shaltiel (University of Haifa, IL), Jacobo Torán (Universität Ulm, DE)
September 11–16, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22371

22372 – Knowledge Graphs and Their Role in the Knowledge Engineering of the 21st Century
Paul Groth (University of Amsterdam, NL), Elena Simperl (King's College London, GB), Marieke van Erp (KNAW Humanities Cluster – Amsterdam, NL), Denny Vrandecic (Wikimedia – San Francisco, US)
September 11–14, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22372

22381 – Rational Design of RiboNucleic Acids
Sven Findeiß (Universität Leipzig, DE), Christoph Flamm (Universität Wien, AT), Yann Ponty (Ecole Polytechnique – Palaiseau, FR)
September 18–23, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22381

22382 – Machine Learning for Science: Bridging Data-Driven and Mechanistic Modelling
Philipp Berens (Universität Tübingen, DE), Kyle Cranmer (University of Wisconsin – Madison, US), Neil D. Lawrence (University of Cambridge, GB), Jessica Montgomery (University of Cambridge, GB), Ulrike von Luxburg (Universität Tübingen, DE)
September 18–23, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22382

22391 – Cognitive Robotics
Fredrik Heintz (Linköping University, SE), Gerhard Lakemeyer (RWTH Aachen, DE), Sheila McIlraith (University of Toronto, CA)
September 25–30, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22391

22392 – Transparent Quantitative Research as a User Interface Problem
Kasper Hornbaek (University of Copenhagen, DK), Yvonne Jansen (CNRS – Talence, FR), Amelia A. McNamara (University of St. Thomas – St. Paul, US), Judy Robertson (University of Edinburgh, GB), Chat Wacharamanotham (Swansea University, GB)
September 25–30, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22392

22401 – Computer Science Methods for Effective and Sustainable Simulation Studies
Wentong Cai (Nanyang TU – Singapore, SG), Christopher Carothers (Rensselaer Polytechnic Institute – Troy, US), David M. Nicol (University of Illinois – Urbana Champaign, US), Adelinde M. Uhrmacher (Universität Rostock, DE)
October 3–7, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22401

22402 – Foundations for a New Perspective of Understanding Programming
André Brechmann (Leibniz-Institut für Neurobiologie – Magdeburg, DE), Bonita Sharif (University of Nebraska – Lincoln, US), Janet Siegmund (TU Chemnitz, DE), Westley Weimer (University of Michigan – Ann Arbor, US)
October 3–7, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22402

22411 – Theory and Practice of SAT and Combinatorial Solving
Olaf Beyersdorff (Friedrich-Schiller-Universität Jena, DE), Armin Biere (Universität Freiburg, DE), Vijay Ganesh (University of Waterloo, CA), Jakob Nordström (University of Copenhagen, DK & Lund University, SE)
October 9–14, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22411

22412 – Intelligent Security: Is "AI for Cybersecurity" a Blessing or a Curse
Lejla Batina (Radboud University Nijmegen, NL), Annelie Heuser (CNRS – IRISA – Rennes, FR), Nele Mentens (Leiden University, NL), Stjepan Picek (TU Delft, NL), Ahmad-Reza Sadeghi (TU Darmstadt, DE)
October 9–14, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22412

22421 – Security of Decentralized Financial Technologies
Joseph Bonneau (New York University, US), Arthur Gervais (Imperial College London, GB), Marie Vasek (University College London, GB)
October 16–21, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22421

22422 – Developmental Machine Learning: From Human Learning to Machines and Back
Pierre-Yves Oudeyer (INRIA – Bordeaux, FR), James M. Rehg (Georgia Institute of Technology – Atlanta, US), Linda B. Smith (Indiana University – Bloomington, US), Sho Tsuji (University of Tokyo, JP)
October 16–21, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22422

22431 – Data-Driven Combinatorial Optimisation
Emma Frejinger (University of Montreal, CA), Andrea Lodi (Cornell Tech – New York, US), Michele Lombardi (University of Bologna, IT), Neil Yorke-Smith (TU Delft, NL)
October 23–28, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22431

22432 – Towards a Unified Model of Scholarly Argumentation
Khalid Al-Khatib (University of Groningen, NL), Anita de Waard (Elsevier – Jericho, US), Iryna Gurevych (TU Darmstadt, DE), Yufang Hou (IBM Research – Dublin, IE)
October 23–28, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22432

22441 – Optimization at the Second Level
Luce Brotcorne (INRIA Lille, FR), Christoph Buchheim (TU Dortmund, DE), Dick den Hertog (University of Amsterdam, NL), Gerhard J. Woeginger (in memoriam; † April 1, 2022)
October 30 to November 4, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22441

22442 – Toward Scientific Evidence Standards in Empirical Computer Science
Brett A. Becker (University College Dublin, IE), Christopher D. Hundhausen (Oregon State University – Corvallis, US), Ciera Jaspan (Google – Mountain View, US), Andreas Stefik (University of Nevada – Las Vegas, US), Thomas Zimmermann (Microsoft Corporation – Redmond, US)
October 30 to November 4, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22442

22451 – Principles of Contract Languages
Dilian Gurov (KTH Royal Institute of Technology – Stockholm, SE), Reiner Hähnle (TU Darmstadt, DE), Marieke Huisman (University of Twente – Enschede, NL), Giles Reger (University of Manchester, GB)
November 6–11, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22451

22452 – Computational Social Dynamics
Martin Hoefer (Goethe-Universität – Frankfurt am Main, DE), Sigal Oren (Ben Gurion University – Beer Sheva, IL), Roger Wattenhofer (ETH Zürich, CH)
November 6–11, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22452

22461 – Dynamic Graph Algorithms
Aaron Bernstein (Rutgers University – New Brunswick, US), Shiri Chechik (Tel Aviv University, IL), Sebastian Forster (Universität Salzburg, AT), Tsvi Kopelowitz (Bar-Ilan University – Ramat Gan, IL)
November 13–18, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22461

22462 – Set Visualization and Uncertainty
Susanne Bleisch (FH Nordwestschweiz – Muttenz, CH), Steven Chaplick (Maastricht University, NL), Jan-Henrik Haunert (Universität Bonn, DE), Eva Mayr (Donau-Universität Krems, AT), Marc van Kreveld (Utrecht University, NL)
November 13–18, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22462

22471 – Towards More Flexible and Automated Communication Networks
Rui Aguiar (Institute of Telecommunications – Aveiro, PT), Artur Hecker (Huawei Technologies – München, DE), Stefan Schmid (TU Berlin, DE), Henning Schulzrinne (Columbia University – New York, US)
November 20–23, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22471

22481 – Vertex Partitioning in Graphs: From Structure to Algorithms
Maria Chudnovsky (Princeton University, US), Neeldhara Misra (Indian Institute of Techology – Madras, IN), Daniel Paulusma (Durham University, GB), Oliver Schaudt (Bayer AG – Leverkusen, DE)
November 27 to December 2, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22481

22482 – Counting and Sampling: Algorithms and Complexity
Holger Dell (Goethe-Universität – Frankfurt am Main, DE), Mark R. Jerrum (Queen Mary University of London, GB), Haiko Müller (University of Leeds, GB)
November 27 to December 2, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22482

22491 – Cognitive Augmentation
Kai Kunze (Keio University – Yokohama, JP), Pattie Maes (MIT – Cambridge, US), Florian 'Floyd' Mueller (Monash University – Clayton, AU), Katrin Wolf (Berliner Hochschule für Technik, DE)
December 4–9, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22491

22492 – Formal Methods and Distributed Computing: Stronger Together
Hagit Attiya (Technion – Haifa, IL), Constantin Enea (Ecole Polytechnique – Palaiseau, FR & CNRS – Palaiseau, FR), Sergio Rajsbaum (National Autonomous University of Mexico, MX), Ana Sokolova (Universität Salzburg, AT)
December 4–9, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22492

22512 – Inter-Vehicular Communication – From Edge Support to Vulnerable Road Users II
Ana Aguiar (Universidade do Porto, PT), Onur Altintas (Toyota Motors North America – Mountain View, US), Falko Dressler (TU Berlin, DE), Gunnar Karlsson (KTH Royal Institute of Technology – Stockholm, SE)
December 18–21, 2022 | Dagstuhl Seminar | https://www.dagstuhl.de/22512

Dagstuhl-Perspektiven-Workshops

14.2

Dagstuhl Perspectives Workshops

22262 – Human-Centered Artificial Intelligence
Virginia Dignum (University of Umeå, SE), Wendy E. Mackay (INRIA Saclay – Orsay, FR), John Shawe-Taylor (University College London, GB), Frank van Harmelen (VU University Amsterdam, NL)
June 26 to July 1, 2022 | Dagstuhl Perspectives Workshop | https://www.dagstuhl.de/22262

22282 – Current and Future Challenges in Knowledge Representation and Reasoning
James P. Delgrande (Simon Fraser University – Burnaby, CA), Birte Glimm (Universität Ulm, DE), Thomas Meyer (University of Cape Town, ZA), Miroslaw Truszczynski (University of Kentucky – Lexington, US), Frank Wolter (University of Liverpool, GB)
July 10–15, 2022 | Dagstuhl Perspectives Workshop | https://www.dagstuhl.de/22282

GI-Dagstuhl-Seminare 14.3 *GI-Dagstuhl Seminars*

There were no such meetings in 2022.

Lehrveranstaltungen 14.4 *Educational Events*

22253 – Summer School "Data Management Techniques"
Goetz Graefe (Google – Madison, US)
June 19–23, 2022 | Educational Event | https://www.dagstuhl.de/22253

22503 – Lehrerfortbildung in Informatik
Sebastian Connette (Landesinstitut f. Pädagogik & Medien – Saarbrücken, DE), Michael Gerke (Schloss Dagstuhl – Saarbrücken, DE), Hannes Heusel (Eduard-Spranger Gymnasium Landau, DE & Pädagogisches Landesinstitut Rheinland-Pfalz, DE)
December 14–16, 2022 | Educational Event | https://www.dagstuhl.de/22503

22504 – inf-schule.de: Weiterentwicklung des elektron. und interakt. Lehrbuches für den Informatikunterricht
Hannes Heusel (Eduard-Spranger Gymnasium Landau & Pädagogisches Landesinstitut Rheinland-Pfalz, DE)
December 13–14, 2022 | Educational Event | https://www.dagstuhl.de/22504

Forschungsgruppentreffen 14.5 *Research Group Meetings*

22013 – Decision-Making under Uncertainty – From Theory to Practice
Nils Jansen (Radboud University Nijmegen, NL)
January 3–7, 2022 | Research Group Meeting | https://www.dagstuhl.de/22013

22023 – Doktorandenseminar FH Aachen
Philipp Kohl (FH Aachen – Jülich, DE), Bodo Kraft (FH Aachen – Jülich, DE)
January 11–14, 2022 | Research Group Meeting | https://www.dagstuhl.de/22023

22073 – Complexity of Infinite-Dimensional Problems
Klaus Ritter (TU Kaiserslautern, DE)
February 13–18, 2022 | Research Group Meeting | https://www.dagstuhl.de/22073

22093 – Modeling and Solutions for Decision-Making in Smart Semiconductor Supply Chain and Supply Chains containing Semiconductors
Hans Ehm (Infineon Technologies – München, DE)
February 27 to March 4, 2022 | Research Group Meeting | https://www.dagstuhl.de/22093

22115 – Arbeitskreis KI in Schulen
Ute Schmid (Universität Bamberg, DE)
March 13–16, 2022 | Research Group Meeting | https://www.dagstuhl.de/22115

22116 – Klausurtagung Telematik
Martina Zitterbart (KIT – Karlsruher Institut für Technologie, DE)
March 16–18, 2022 | Research Group Meeting | https://www.dagstuhl.de/22116

22119 – Forschungsaufenthalt "CERG AI education initiative"
Ralf Romeike (Freie Universität Berlin, DE)
March 16–18, 2022 | Research Group Meeting | https://www.dagstuhl.de/22119

22123 – OpenML meeting
Joaquin Vanschoren (TU Eindhoven, NL)
March 20–25, 2022 | Research Group Meeting | https://www.dagstuhl.de/22123

22133 – Software Engineering Forschungsmethoden Training
Lars Grunske (HU Berlin, DE)
March 30 to April 1, 2022 | Research Group Meeting | https://www.dagstuhl.de/22133

22134 – Klausurtagung der Arbeitsgruppe "Algorithmische Grundlagen"
Johannes Fischer (TU Dortmund, DE)
March 27–30, 2022 | Research Group Meeting | https://www.dagstuhl.de/22134

22153 – GIBU 2022: GI-Beirat der Universitätsprofessor*innen
Lars Grunske (HU Berlin, DE)
April 10–12, 2022 | Research Group Meeting | https://www.dagstuhl.de/22153

22154 – Probabilistic Verification
Thomas Noll (RWTH Aachen University, DE)
April 10–14, 2022 | Research Group Meeting | https://www.dagstuhl.de/22154

22173 – Applied Machine Intelligence
Johannes Busse (HAW Landshut, DE), Thomas Hoppe (Fraunhofer FOKUS – Berlin, DE), Bernhard Humm (Hochschule Darmstadt, DE), Anatol Reibold (Cogia GmbH – Frankfurt am Main, DE)
April 27–29, 2022 | Research Group Meeting | https://www.dagstuhl.de/22173

22209 – Squares in Permutations
Guillaume Fertin (Nantes Université, FR)
May 15–20, 2022 | Research Group Meeting | https://www.dagstuhl.de/22209

22213 – Kolloquium zum GI Dissertationspreis 2021
Steffen Hölldobler (TU Dresden, DE)
May 22–25, 2022 | Research Group Meeting | https://www.dagstuhl.de/22213

22214 – DFG Rundgespräch
Susanne Boll (Universität Oldenburg, DE)
May 22–25, 2022 | Research Group Meeting | https://www.dagstuhl.de/22214

22243 – Gemeinsamer Workshop der Graduiertenkollegs
Felix Freiling (Universität Erlangen-Nürnberg, DE), Helmut Seidl (TU München, DE)
June 12–15, 2022 | Research Group Meeting | https://www.dagstuhl.de/22243

22244 – Informatik und Recht – Synergien für den Fortschritt
Maximilian Eichacker (Universität des Saarlandes, DE)
June 12–15, 2022 | Research Group Meeting | https://www.dagstuhl.de/22244

22273 – Dissertation Writing Workshop
Felix Schwinger (RWTH Aachen, DE)
July 3–8, 2022 | Research Group Meeting | https://www.dagstuhl.de/22273

22353 – Interoperability for Cross-Domain Research: Machine-Actionability & Scalability
Simon Cox (CSIRO – Clayton South, AU), Arofan Gregory (DDI Alliance / CODATA, US), Simon Hodson (CODATA – Paris, FR), Steven McEachern (Australian National University – Acton, AU), Hilde Orten (SIKT – Bergen, NO), Joachim Wackerow (Ludwigshafen, DE)
August 28 to September 2, 2022 | Research Group Meeting | https://www.dagstuhl.de/22353

22473 – Klausurtagung der Promovierenden des FB Informatik, TU Kaiserslautern
Annette Bieniusa (TU Kaiserslautern, DE)
November 24–25, 2022 | Research Group Meeting | https://www.dagstuhl.de/22473

22474 – European Laboratory's for Learning and Intelligent Systems Third ELLIS NLP Workshop
Iryna Gurevych (TU Darmstadt, DE), André F. T. Martins (IST – Lisbon, PT), Victor Prokhorov (University of Edinburgh, GB), Ivan Titov (University of Edinburgh, GB), Xinnuo Xu (University of Edinburgh, GB)
November 20–23, 2022 | Research Group Meeting | https://www.dagstuhl.de/22474

22475 – Workshop "Language-Based Expert-AI Cooperation" (LExIC) of TU Darmstadt and LMU
Maral Dadvar (TU Darmstadt, DE), Iryna Gurevych (TU Darmstadt, DE), Hinrich Schütze (LMU München, DE)
November 23–25, 2022 | Research Group Meeting | https://www.dagstuhl.de/22475

22476 – Fuzzing! Languages! Reasoning! Learning! Group Retreat Zeller 2022
Andreas Zeller (CISPA – Saarbrücken, DE)
November 23–25, 2022 | Research Group Meeting | https://www.dagstuhl.de/22476

22513 – Doktorandenseminar "Software-Engineering in der Robotik", Abt. Robotik, Fraunhofer IPA

Björn Kahl (Fraunhofer IPA – Stuttgart, DE)

December 18–21, 2022 | Research Group Meeting | https://www.dagstuhl.de/22513

Leibniz
Gemeinschaft

Oktavie-Allee, 66687 Wadern, Deutschland

Jahresbericht / Annual Report 2022 | ISSN 2199-1995
http://www.dagstuhl.de