Report from Dagstuhl Seminar 18412

Encouraging Reproducibility in Scientific Research of the Internet

Edited by

Vaibhav Bajpai[1], Olivier Bonaventure[2], Kimberly Claffy[3], and Daniel Karrenberg[4]

1 **TU München, DE,** bajpaiv@in.tum.de
2 **UC Louvain, BE,** olivier.bonaventure@uclouvain.be
3 **San Diego Supercomputer Center, US,** kc@caida.org
4 **RIPE NCC - Amsterdam, NL,** dfk@ripe.net

Abstract

Reproducibility of research in Computer Science (CS) and in the field of networking in particular is a well-recognized problem. For several reasons, including the sensitive and/or proprietary nature of some Internet measurements, the networking research community pays limited attention to the of reproducibility of results, instead tending to accept papers that appear plausible.

This article summarises a 2.5 day long Dagstuhl seminar on Encouraging Reproducibility in Scientific Research of the Internet held in October 2018. The seminar discussed challenges to improving reproducibility of scientific Internet research, and developed a set of recommendations that we as a community can undertake to initiate a cultural change toward reproducibility of our work. It brought together people both from academia and industry to set expectations and formulate concrete recommendations for reproducible research. This iteration of the seminar was scoped to computer networking research, although the outcomes are likely relevant for a broader audience from multiple interdisciplinary fields.

Seminar October 7–10, 2018 – http://www.dagstuhl.de/18412
2012 ACM Subject Classification Networks
Keywords and phrases Computer Networks, Reproducibility
Digital Object Identifier 10.4230/DagRep.8.10.41

1 Executive Summary

Vaibhav Bajpai
Olivier Bonaventure
Kimberly Claffy
Daniel Karrenberg

Reproducibility in scientific research is a means to not only achieve trustworthiness of results, but it also lowers barriers to technology transition [40] and accelerates science by promoting incentives to data sharing. The networking research community however pays limited attention to the importance of reproducibility of results, instead tending to accept papers that appear plausible. Previous studies [29, 41, 18] have shown that a fraction of published papers release artifacts (such as code and datasets) that are needed to reproduce results. In order to encourage reproducibility of research, practitioners continue [33, 28, 37, 11, 20] to do service to educate the community on the need for this change. To provide incentives to authors, vehicles for publication of software and datasets are also emerging. For instance, Elsevier SoftwareX [3] is a new journal designed to specifically publish software contributions. DataCite [36, 27] provides mechanisms for supporting methods to locate and cite datasets. Community Resource for Archiving Wireless Data (CRAWDAD) [43] and Information Marketplace for Policy and Analysis of Cyber-risk & Trust (IMPACT) Cyber Trust [4] provide an index of existing measurement data to not only enable new research but also advance network science by promoting reproducible research. Traditional conferences bestow best dataset awards and actively solicit submissions that reproduce results. SIGCOMM Computer Communication Review (CCR) allows authors to upload artifacts during paper submission to allow reviewers to check for reproducibility, and relaxes page limits for reproducible papers. Association for Computing Machinery (ACM) has recently introduced a new policy [1] on result and artifact review and badging. The policy identifies a terminology to use to assess results and artifacts. ACM has also initiated a new task force on data, software and reproducibility in publication [8] to understand how ACM can effectively promote reproducibility within the computing research community. National Academies of Sciences, Engineering, and Medicine with a goal to move towards the open science ecosystem has recently (2018) released a report [31] with guidance and concrete recommendations on how to build strategies for achieving open science. The target is to ensure the free availability (and usability) of publications and associated artifacts. The National Science Foundation (NSF) is taking substantial steps [2] in this area whereby submitted proposals are required to provide a results dissemination plan to describe how produced research results are made available to the extent necessary to independently validate the findings. Towards this end, the proposal budget [5] may request funds for the costs of documenting, preparing, publishing or otherwise making available to others the findings and products of the work conducted under the NSF grant. Despite these continued efforts, reproducibility of research exist as an ongoing problem and few papers that reproduce existing research get published [17, 26, 34] in practise.

Goals

In this seminar, we discussed challenges to improving reproducibility of scientific Internet research, developed a set of recommendations that we as a community can undertake to initiate a cultural change toward increased reproducibility of our work. The goal of the seminar was to discuss the questions below and to propose recommendations that would improve the state of reproducibility in computer networking research.

- **What are the challenges with reproducibility?**

  How can researchers (and data providers) navigate concerns with openly sharing datasets? How should we cope with datasets that lack stable ground truth?

The first category of questions tried to identify the challenges with reproducibility [14]. For instance, concerns with openly sharing datasets led to discussions around legal restrictions and the advantages of researchers keeping data private for their own exclusive future use. Another consideration is double-blind review practices, which require that authors expend effort to obfuscate the source of their data. Would this time be better spent documenting the datasets for sharing to enable reproducibility? A “gap analysis” discussion to understand whether the problem is a lack of appropriate venues or lack of stable ground truth, or more broadly a lack of incentive to reproduce research since publishing (and funding) agents tend to prefer novelty was held. There is also the inherent risk of confirmation bias of existing results; discussion of ideas on how to train young researchers to recognize and counter this tendency was sought.

- **What incentives are needed to encourage reproducibility?**

  What can publishers do? What can conference organisation committees do? How can we ensure that reviewers consider reproducibility when reviewing papers? How can we manage and scale the evaluation of artifacts during peer review? Do we need new venues that specifically require reproducibility of the submitted research?

The second category of questions is about incentives. Questions about how publishers can promote reproducibility framed discussions on whether publishers can provide storage for authors to upload data artifacts with the associated paper in digital libraries, or whether mechanisms can be developed to highlight reproducible (and reproduced) papers. Questions on how conference organisation committees can inspire ideas for additional incentives (such as best dataset awards or relaxing page limits) for authors to make research reproducible. We identified questions to add to review forms to ensure reviewers pay attention to reproducibility aspects. This further lead to discussions on whether committees (in parallel to the regular technical program committee) should evaluate artifacts during the conference review process. Should such a committee be composed of purely young researchers or a blend of young and senior researchers? Questions on the need for specific venues triggered discussions on whether high-impact journals need to establish feature topics on reproducibility or devote a dedicated column for papers that reproduce existing research.

- **What tools and systems are available to facilitate reproducibility?**

  How effective are emerging interactive lab notebook tools (e.g., Jupyter) at facilitating reproducibility? Should CS course curricula integrate use of these tools for student projects to help develop skills and habits that enable reproducibility?

The third category of questions attempt to identify and review tools and systems that are available to facilitate reproducibility. Enormous interest has developed recently in tools for recording experimental observations and computational analytics on large data sets. Some researchers now document the entire process for a paper in a Jupyter lab notebook, greatly facilitating reproducibility and extension of the research. The learning curve for these tools may be daunting; we discussed how faculty can evolve CS course curricula to integrate use of these tools for student projects to help develop skills and habits that enable reproducibility.

- **What guidelines or (best practises) are needed to help reproducibility?**

  How can we ensure authors think about reproducibility? What guidelines would assist reviewers in evaluating artifacts?

The fourth category of questions attempts to develop guidelines (or best practises) to promote reproducibility of research. For instance, we discussed what language could be added to Call for Papers (CFP) to encourage authors to describe reproducibility aspects (of both measurements and results) in their paper submissions.

Structure

The seminar lasted 2.5 days. The seminar began with an introductory round where each participant presented one slide to give an overview of their experience that is relevant for the seminar and a set of open questions that the participant wished to discuss during the event. These slides were collected from each participant before the seminar. We had one invited talk (§3.1) that we used as a basis for triggering discussions and identifying areas for group work, while a major portion of the seminar time was dedicated to breakout sessions, whereby participants were split into small groups to discuss specific themes and develop ideas with consensus to propose to larger groups. The morning sessions the following day were dedicated to continuing parallel group work with presentations that reported the outcomes of each breakout session from the previous day. In the afternoons, we dedicated some time for seven minute lightning talks to invite ideas for subsequent breakout sessions. One evening, we had a social dinner activity. The afternoon of the third day was spent reviewing and collecting feedback from the participants and to initiating follow up actions identified during the seminar.

2 Table of Contents

3 Presentations

Participants were encouraged to volunteer for a lightning talk to provide their perspective on the topic and the presentations were intended as a basis for triggering discussions and identifying areas for breakout sessions.

3.1 Hyper papers and Open Co-Authoring

Alberto Dainotti (CAIDA) kicked off the discussion by presenting some history of the scientific publication process. Scientific papers were born as a mean to share novel scientific knowledge. However, over time publications have also become the main metric for career advancement. This shift has influenced the whole publishing process, from the generation of ideas, data and results to how they are shared. He proposed that perhaps there is a need to step back and look at the currently established process for scientific paper authoring and publishing, including conventions and formats, and wondered if there were room for optimization for the good of science and education. For example, have we struck the right balance between "secrecy" and openness? Are there opportunities from recent technologies and collaborative practices that we can leverage to address the following relevant issues: (i) ideas are often kept secret until a paper is published; (ii) studies are often not reproducible; (iii) incremental work is discouraged by lack of incentives and practical barriers; (iv) fixed-layout flat documents have limitations which are not addressed by simply attaching supplemental material. He proposed to explore the concept of "open collaborative hyperpapers": a paper writing paradigm where co-authorship is potentially open to any researcher, using formats and tools that by design incorporate reproducibility and accountability of contributions, enable incremental progress, and allow for experimenting with different models of code/data/paper reviewing. The talk led to a parallel group breakout (§4.1) where the idea was further developed.

3.2 SIGCOMM Reproducibility Workshop and Artifacts Survey

Recent years have shown an increasing awareness of issues of reproducibility of results as an essential part of research. To address this important issue, ACM has introduced a new policy [1] on result and artifacts review and badging. The policy defines the terminology to be used to assess results and artifacts but does not specify the review process or how to make research reproducible. Furthermore, there appears to be an inconsistency with the terminology defined by other scientific communities [7]. This concern led to a parallel group breakout (§4.5) where specifically the ontology of reproducibility was formalised. At SIGCOMM 2017, a workshop was organized with the specific purpose to tackle the issue. The objective was to trigger discussion and activity in order to craft recommendations on how to introduce incentives for authors to share their artifacts, and the details on how to use them, as well as defining the process to be used to evaluate reproducibility. Luigi Iannone (Télécom ParisTech) presented an overview of this workshop [38] and summarized the main discussions and outcomes.

To improve the current state and strengthen ongoing community efforts towards reproducibility, as a followup to the SIGCOMM reproducibility workshop, a survey was conducted among the authors of papers published at leading ACM computer networking conferences in 2017: CoNEXT, ICN, IMC, and SIGCOMM. Damien Saucez (INRIA Sophia Antipolis) presented the current state of artifacts availability and reproducibility based on this survey [23]. The goal of the survey was to serve as a starting point for further discussions to encourage researchers to ease the reproduction of scientific work published within the SIGCOMM community and to inspire program chairs of future conferences to emphasize the importance of reproducible research.

3.3 Artifacts Evaluation Committee & CoNEXT'18 Badges

Damien Saucez (INRIA Sophia Antipolis) introduced the new ACM SIGCOMM Artifacts Evaluation Committees (AEC) (similar to the AEC created in several SIGs or conferences). The objective of the AEC is to evaluate the artifacts of papers accepted at the SIGCOMM sponsored conferences and assign badges (using the badging system [1] recently established by the ACM) to these papers. Every paper of six pages or more that has been published or accepted by SIGCOMM CCR or any of the conferences sponsored by ACM SIGCOMM in 2018 was eligible for artifacts evaluation by the ACM SIGCOMM AEC. Authors submitted a revised version of their accepted paper that also included pointers to their publicly-available artifacts and in the appendix additional information to help the reviewers with evaluation of the artifact. The ACM Digital Library (DL) was updated to attach assigned badges and public reviews to all the badged papers. As of December 2018, 37 volunteers evaluated 32 submitted papers as part of the AEC[1]. Meanwhile, papers recently accepted at CoNEXT 2018 underwent such a badging process. 14 papers out of 32 accepted papers at CoNEXT were evaluated and 12 were awarded with badges with the help of 20 volunteers who evaluated the artifacts of submitted papers. The lessons learned from the evaluation of the these submitted artifacts will be appear as an editorial note that will be published in the beginning of 2019, with all public reviews and the list of badged papers.

3.4 Challenges with Reproducibility

Mirja Kühlewind (ETH Zürich) raised concerns on how the CS culture is receptive to accepting papers that are non-reproducible so long as they appear plausible. In her talk [14], she discussed some of the challenges that hinders authors and reviewers to embrase reproducibility. For instance, lack of dedicated publishing venues reduce the incentives for authors to reproduce existing research. When submitting papers to double-blind venues, authors have to spend time obfuscating the manuscript, which could instead be used to make artifacts available. On the other hand, paper submission systems do not generally allow authors to upload artifacts, which compels reviewers to fetch artifacts from provided locations discounting their anonymity during the review process. There is also lack of incentive to commit significant time to artifact review to ensure reproducibility of research.

Towards this end, she proposed a set of recommendations. For instance, authors can be encouraged to discuss reproducibility considerations in papers. Paper submission systems can provide authors mechanisms to upload artifacts for reviewing purposes and the review forms themselves can be augmented to guide reviewers to think about reproducibility of the submitted manuscript. After acceptance, reproducible papers can be also be highlighted in digital libraries for recognition purposes.

[1] https://sigcomm18ae.hotcrp.com

3.5 Experiences on Reproducing a Routing Security Paper

In this talk, Matthias Wählisch (FU Berlin) reported about his experiences with reproducing a paper about routing security [35]. He identified that data sources that are used to analyze secure inter-domain routing, usually lack sufficient description. He illustrated that the selection of vantage points is crucial within this context. Based on his experiences, he concluded with three observations. First, many authors are afraid to share their tools because they are afraid to reveal mistakes. Second, asking for reproducibility is especially important in inter-disciplinary research as this allows to self-check the level of competence of the reviewers. Finally, the community needs a change in culture: Making mistakes is not preferred but denying mistakes is worse. Resolving these pieces might help to advance reproducible research.

3.6 Reproducibility: A Problem of Economics, Not Science

Henning Schulzrinne (Columbia University) identified that lack of replicability (see [1] for a definition) tends to pollute the knowledge pool. He argued that reproducibility is a matter of aligning incentives, and the incentives all argue against reproducibility. Replicability has a higher opportunity cost, and is associated with a principal-agent problem where the funding bodies may want to encourage replication, but the researchers may not due to lack of incentives. This talk led to a parallel group breakout (§4.4) where the issue of aligning incentives for reproducibility was further discussed.

3.7 (Strict) Reproducibility Considered both Hard and Harmful

Much current discussion of reproducibility focuses on reproduction of the specific, concrete result obtained by the original researcher. John Wroclawski (USC) proposed to focus instead on reproduction of results at the semantic level, with the aim of validating the larger research conclusion of the original work rather than obtaining a precisely identical narrow result. To support this goal, John suggested the use of explicitly specified invariants to bound the region of applicability of a result, and then briefly touch on the range of discovery, specification, and enforcement tools that, if they existed, would facilitate the use of invariants in support of reproducible research. This talk led to a parallel group breakout (§4.5) where the ontology of reproducibility was formalised.

3.8 Interactive Data as a New Publication Model for Journals

Ralph Holz (The University of Sydney) proposed to make changes to the publication model of journals: instead of producing, and being the gatekeeper to the equivalent of a printout, publications should be ’containerised’ - like websites, they should be runnable applications, with the real dataset in the background, where users can choose the appropriate form of presentation and even apply filters and make changes to code. The new form of publication combines write-up, dataset, and software in an instance that is playable and reproducible and makes reviews much easier and tractable. He discussed both advantages and possible problems that went into further discussion into the parallel group (§4.1) that converged into the concept of hyperpapers and new publication strategies.

3.9 Reproducibility vs. Measurement Infrastructures

Robert Kisteleki (RIPE NCC) highlighted a few examples of reproducibility issues related to measurement frameworks [15] such as (geographical and topological) biases, (stability) of vantage point allocation [13, 25], timing of the research, and unique properties of measurement infrastructures in relation to capabilities, data access formats, and data anonymisation. This talk led to the development of a parallel group that discussed data access formats (§4.8) in more detail.

3.10 Taming the Complexity of Artifact Reproducibility

Reproducing research results, as it is required for peer review, can be a time-consuming and difficult task. Thomas Zinner (TU Berlin), proposed three approaches to improve how research results can be substantiated and discussed their applicability. The proposals are based on a brief study [22] on evaluation methods (for Software-defined Network (SDN) research) and insights from a comprehensive discussion on reproducibility. The first approach proposes the use of 'meta-artifacts', which he defined as a structured piece of metadata that describes the tools and parameters that are used during the evaluation. He envisioned a community driven database holding a well-functional set of such meta-artifact templates that could assist in the documentation of the evaluation process. The second approach proposes to either share domain-specific 'evaluation environments' or at least establish well-known 'evaluation scenarios'. For instance, a description of traffic patterns and topology to realise a representative campus network could be one such evaluation scenario. The third approach proposes to adopt 'self-provisioning evaluation setups' using Vagrant (or Docker) to help reproduce the circumstances as close as possible to the original experiment. Some of these ideas became input for the parallel group breakout (§4.1) on new publication strategies.

3.11 Towards an Ecosystem of Reproducibility

Quirin Scheitle (TU München) argued that changing the culture in the network measurement field towards more reproducible results requires changes at many elements of the ecosystem, including authors and independent reproducers. The talk explored what these incentives [39] might be and what might be key factors to their success. Examples to get more reproductions may be making them a soft requirement in PhD programs, co-locating replication hackathons at major venues, or turning labs at PhD schools [6] into replication efforts. Some of these ideas were discussed further in the parallel group (§4.4) on creating incentives of reproducibility.

3.12 High-Quality Measurements and Modelling of Packet Processing

There is a trend towards increasing complexity of networked systems, which leads to the challenge to produce high-quality data for reproducible research. Georg Carle (TU München) presented methods to measure in a reproducible manner and to assess the quality of results. He presented four dimensions to assess the quality of measurement data based on precision, accuracy, coverage and scope. Precision and accuracy requires high-quality tools with hardware support for time-stamping and rate control, thereby limiting random and systematic errors of the traffic generation process. With respect to hard real-time guarantees [24], measurement data with high coverage requires various types of artificial load for rare and worst-case system states. Measurement data of higher “scope” characterises widely used hardware/software configurations. The tool MoonGen [21], a dedicated packet generator in broad use by the community, with its hardware-supported measurement capabilities supports these dimensions of high-quality data, thereby allowing network experiments to reliably reproduce measurement results.

3.13 Measuring Mobile Broadband Networks with MONROE

MONROE [9] is a platform for mobile broadband measurements explicitly designed with openness and reproducibility in mind. In this talk, Anna Brunstrom (Karlstad University) introduced MONROE and discussed its design from a reproducibility and repeatability perspective. MONROE measurement nodes are deployed both in fixed locations and on board trains and buses, offering the possibility to measure from the same vantage points or along the same routes over time. The platform is open to external users, allowing researchers to repeat or extend previous measurements. It is built on open source software and open hardware specifications, allowing others to extend the platform or reuse its design. Experiments in MONROE are designed as Docker containers, making them easy to reuse by others in the same or other environments. Rich metadata is available on the platform and saved in the MONROE database to provide context for the measurements. Open data is made available for all results published by the MONROE alliance.

3.14 Reproducible Research: Implications of Roaming in Europe

“Roam like Home” is the initiative of the European Commission (EC) to end the levy of extra charges when roaming within the European region. As a result, people are able to use data services more freely across Europe. This brings the need for operators to provide seamless service for their customers, similar to what they experience in their home country. However, the implications roaming solutions have on performance have not been carefully examined. In this talk, Andra Lutu (Telefónica Research) presented how they leveraged MONROE open source components to build a roaming measurement infrastructure [30] for 16 different mobile networks in six countries across Europe. With this infrastructure, they then measured different aspects of international roaming in 3G and 4G networks, including mobile network configuration, performance characteristics, and content discrimination. To facilitate reproducibility, Andra plans to extend this research into a hyperpaper which was further developed in the parallel group (§4.1) on new publication strategies.

3.15 Observatories for Internet Measurement

Brian Trammell (ETH Zürich) presented a general model for “measurement observatories”, which supports comparability, repeatability, and protection of raw measurement data. This model is based around a metadata-first workflow, and normalizers that translate heterogeneous raw data into a common observation schema defined for a given measurement purpose. Metadata attached to raw data, normalizers, analyzers, observations, and queries allow tracking the provenance of each object. This model is validated through the implementation of the observatory for Internet path transparency measurements. The ideas presented in this talk were further developed in the parallel group (§4.8) on data formats.

4 Parallel Group Work

The afternoon sessions were used to discuss certain topics in more depth in smaller groups. This section summarises the discussions of each group.

4.1 New Publication Strategies

Alberto Dainotti, Ralph Holz, Mirja Kühlewind, Andra Lutu, Joel Sommers, Brian Trammell

The group envisioned a publishing ecosystem for Internet science, supporting publications that are self-contained, interactive, multi-level, open, and collaborative. The idea leverages on recent developments in platforms and tools for data science and scientific collaboration to build an experimental publishing ecosystem for Internet measurements based on hyperpapers [10], similar to the WholeTale project [16] that also envisions to unite data products with research articles to create "living publications" (or tales).

Hyperpapers are self-contained and interactive. Ideally, a full hyperpaper contains all the data from which results, plots, and conclusions in the paper are drawn, as well as source code implementing the analytic tasks distilling those results from the raw source data. The paper is interactive, allowing both changes to the raw source data and to the analysis code to be reflected in the analytic products in the paper.

Hyperpapers are multi-level. The initial view of a full hyperpaper includes the typical prose of a paper. Analysis products, such as charts and tables, can be expanded to show how they were derived. However, the paper can also be expanded in other ways. A section of prose may be linked to an alternate view, information for an alternate audience, related content, or a drill down on some interesting set of a result.

The perennial problem of setting up environments for data analysis without needing to replicate a full tool chain with dependencies from scratch is largely solved today by virtualisation and containerization tools such as Vagrant and Docker. Problems of scale are addressed by the easy (if sometimes costly) widespread availability of cloud infrastructure from multiple providers. Integration of data analytics with authoring environment interleaving text and interactive visualizations is supported by data analysis notebooks such as JupyterLab and Apache Zeppelin. GitHub has emerged as the de-facto standard for integrating version control of digital artifacts with a collaboration environment, and its model of working is suited to open collaborative papers, which have a fair amount in common with the long-running open source projects GitHub was originally built to support.

The group identified two main gaps in technical infrastructure necessary for a full initial realization of this vision. First, while some research studies can be done with data or models that can easily be stored in an ad-hoc format within the hyperpaper itself, large-scale Internet measurement studies need access to large data sets mediated through some interface. This exists for certain data sources (such as the RIPE Atlas API), but a full realization would require the creation and standardization of interfaces for retrieval of data and metadata for each broad type of measurement activity. Second, the distribution of rendered versions of papers is currently possible for scientific notebook environments, but these render to a webpage that is not necessarily optimized for accessibility. Tooling to render a view of hyperpaper as a PDF according to the required format for a given venue, is necessary to support the full multi-rendering functionality of the vision above.

An editorial note [19] describing this concept with a call for action to publish demonstrations of hyperpapers and make preparations for an experimental hyperpaper platform recently appeared in SIGCOMM CCR.

4.2 Guidelines for Students

Vaibhav Bajpai, Anna Brunstrom, Anja Feldmann, Wolfgang Kellerer, Aiko Pras, Henning Schulzrinne, Georgios Smaragdakis, Matthias Wählisch, Klaus Wehrle

The group developed guidelines meant for researchers and for students working in experimental networking research early in their career, and as a reminder to others. General best practises on problem formulation and design, documentation, experimentation and data collection and data handling were outlined. For instance, it is essential to formulate the hypothesis, design the experiments to validate (or not) the hypothesis, conduct the necessary experiment, and eventually check the validity of the hypothesis. Planning and soliciting early feedback is crucial in such a workflow, whereby visualisations help to convey early results and help identify anomalies that may need further analysis. One-time experiments are prone to bias by transient effects and dynamism of the operational system in itself, which requires to reiterate the experiment to gain confidence in the results. Documenting all steps and observations (similar to the lab notebook approach common in natural sciences) during experimentation is key for repeatability. The gathered artifacts need to be accompanied with metadata to help understand how the data was created, what it contains and how to recreate it. Embracing version control helps identify regressions in code and analysis to help identify the root cause of the anomaly. It's crucial to keep regular backups to ensure data is safely stored. A good strategy is to run a series of small experiments to verify the tools and validate analysis and then scale up. Identify how not to reinvent the wheel, and which tools are readily available for use in the experiment. During the data collection phase, monitoring should be applied to ensure the smooth running of the experiment to avoid network/disk failures, host reboots, overwritten logs that may distort the data gathering process. It's important to respect the privacy constraints of external datasets that are used in the research. Similarly, before making your datasets available, consult others for any privacy concerns that perhaps could be alleviated by data anonymisation. Furthermore, ensure the integrity of the data to account of observation biases and document them together with the released datasets. Consider how the developed code will be licensed and made available, discuss and form agreement with the team, and perhaps also reach out within your organisation to make yourself aware of the guidelines that may be available. The group went further and also developed specific guidelines by research area, particularly for simulation studies, systems prototyping and evaluations, real-world measurements and subjective experiments along with a recommendations on tools that are generally used in these areas. Pointers to research papers that follow similar guidelines were also identified. An editorial note [12] describing the set of guidelines recently appeared in SIGCOMM CCR.

4.3 Guidelines for Reviewers

Olivier Bonaventure, Luigi Iannone, Daniel Karrenberg, Damien Saucez

This group discussed guidelines for reviewing of artifacts. The aim of the review process is to decide whether to award papers with the relevant ACM badges. The primary document is a form asking a number of questions to the reviewers. The purpose of the form is to structure the discussion among the reviewers. Awarding badges will be a decision based on that discussion. The group realised that this is a work in progress and review methods and standards will converge and best practises be established by the community as we go along. The intent is to provide a starting point. The group started from experience with a review in progress for CONEXT 2018 conference, which already used a form. The group re-worked this form moving some yes/no questions to more differentiated scores that enable reviewers to respond in a more differentiated way. The group also discussed examples across the whole range to provide as guidance to the reviewers. In practice, it is expected that such a review happens between submission of camera ready copy and conference presentation. The rationales are that this allows sufficient time available for review, the paper also becomes immutable after the camera-ready, and badging leads to recognition at the conference. The review form developed by the group is made publicly available[2]. The proposal is to utilise this form and process to review artifacts for upcoming conferences.

4.4 Incentives for Reproducibility

Olivier Bonaventure, Ken Calvert, Luigi Iannone, Damien Saucez, Quirin Scheitle, Jürgen Schönwälder

The group attempted to identify incentives for reproducible papers and workflows to evaluate them. The incentives for independent reproductions of published work and identifying venues for publishing papers reproducing research were studied. Perhaps a carrots (reproducers) and sticks (funding agencies) approach is needed in the long-run to establish a feedback loop to initiate a cultural change. In order to foster a positive view of the process, explicit incentives can be established. The incentives need to be viewed for all players involved. For instance, funding agencies require[3] that results created with their funding are open access with artifacts available (barring cases where data cannot be released due to privacy constraints). Conferences can be made to meet certain reproducibility standards for published papers to get support. Students who try to reproduce results can be handed travel grants to attend conferences and meet authors of papers they reproduce. Publishing venues (such as conferences and journals) can facilitate mechanisms to submit artifacts with published papers and can also integrate them early with the submission process. Reproduction of published results can be made a soft requirement of doctoral studies. A special track in

[2] https://goo.gl/JjXgjw
[3] https://goo.gl/P3L33S

a conference can be established where reproduction reports can be published. Dedicated "repathons" can be organised where authors and reproducers can sign up for attendance backed up by travel support by funding agencies. The repathons can be co-located with regular conferences whereby a list of possible papers available for reproduction are announced in advance. Encouraging reproduction allows authors to develop new ideas with the reproducers and create collaboration possibilities. It also allows the community to build upon the author's work and increases their impact. The badging system helps increase the visibility of work further. Artifact evaluators on the other hand get visibility by being part of the conference committee. The process helps them develop new skills and understanding of the scientific review process.

4.5 Ontology for Reproducibility

Steve Bauer, Georg Carle, Robert Kisteleki, Matt Mathis, Jörg Ott, Karen Sollins, John Wroclawski, Thomas Zinner

The group met to formulate an ontology to be used with reproducibility and their applicability dimensions. The group identified a taxonomy axis composed of observational reproducibility and model (or prediction validation) reproducibility. The former category applies to situations where one is attempting to reproduce a "data collection and analysis" activity. Within this category, the reproducer has no control over the system, but the goal is to collect data with sufficient accuracy to validate the conclusion of an analysis of the data. An example that falls within this category includes real-world measurements on the Internet to understand reality as is, using well-known test-beds. The key reproducibility criteria include vantage point selection, traffic dynamics, methodological description, handling of outliers, and hidden assumptions that need to be documented for increased reproducibility within this category.

The second category applies to situations where one is attempting to reproduce the results of "modeling (or prediction)" activity. The goal is to provide controlled inputs to a system, and then observe how the system responds, so as to observe modeled behavior or validate a predicted result. An example that falls within this category includes simulation experiments that include specific input conditions that do not necessarily arise in the existing system. The key reproducibility criteria include clarity and completeness of description of invariants and dynamic inputs and guidance about the space to explore for predictive study.

Another way to formulate the taxonomy is by identifying the objectives of reproducibility. One objective could be to determine whether a specific result is reproducible, by conceptually keeping all conditions (such as inputs, environment, etc) identical to the original. A different objective could be to reproduce the validity of a result over some range of invariants and inputs. Yet another objective could be on reproducing a methodology to understand whether it is applicable in a certain different circumstance. The key requirement here is to have a precise description of the methodology, along with invariants and some validating test cases.

Yet another way to formulate the taxonomy is by the quality (or strength) of the reproduction activity. For instance, one axes could be the resilience of reproduced result to variations in inputs (or experimental conditions). The precision or degree to which the reproduced result matches the original result could be another axes. The ability to explain (or defend) the proposition of why the reproduced and original results match could be yet another axes. The quality of reproduction is influenced by dynamics of the underlying system. Ideally, reason about appropriate failure to reproduce (vs inappropriate failure to reproduce) in different circumstances should be identified. On the contrary, results that are no longer reproducible for good reason continue to have value for educational use, for re-evaluation with new insights, and for evaluating predictive models.

A completely different view is to ask how effective is a reproducible result as one would go outside from the technical into the non-technical presentation of the result, to the rest of the world. This aspect is particularly important for consumers and policy makers. The clarity of presentation, pre-requisite knowledge and ease of understanding are important factors for considerations to cover this aspect.

4.6 Reproducibility Track

Anna Brunstrom, Georg Carle, Alberto Dainotti, Mirja Kühlewind, Andra Lutu, Damien Saucez, Quirin Scheitle, Georgios Smaragdakis, Matthias Wählisch

The group met with a focus on creating a Call for Papers for a reproducibility track for a conference (or a workshop). The objective of such a reproducibility track is many-fold. For one, the goal is to increase the number of incidents of reproducing a published work. A formal track allows such incidents to be documented in a 2-page reproduction report that can go together with the conference proceedings. The reproduction report shall provide a summary of the original paper, description of the reproduction process and findings including describing challenges with reproducing a certain piece of published result. The authors may be contacted for assistance, but may not become authors of the reproduction report. Existing templates [42] used by universities in their reproducibility seminars can be used as starting point for reproduction reports. The target for a conference (instead of workshop) track allows higher attendance and increases focus on in-depth quality. This also has the side-benefit of giving more visibility to authors that publish reproducible research. Furthermore, the track explicitly allows early-stage researchers to learn not only how to reproduce a published research, but also document and report about it. The track also gives new researchers an opportunity to get in touch or work directly with the potentially more senior authors of the original paper, which can be an additional incentive to participate. Unlike the AEC, where the focus is on replication, the focus of the reproducibility track is on independent reproduction of results by interested students. The process shall also uncover sanity of the employed methodology and eventually recognise the reproducers by recognising their contribution as a citable publication. The AEC badges can be used to identify papers that can be used as input for the reproducibility track. The authors of the original publication shall be allowed to provide a commentary on the reproduction report.

The proposal for a reproducibility track was made at the Internet Measurement Conference (IMC) 2018 conference and is currently under discussion.

4.7 Reproducibility in Post-publication Phase

Olivier Bonaventure, Kimberly Claffy, Luigi Iannone and Henning Schulzrinne

The group met with a goal to understand how reproducibility can be maintained in the post-publication phase. The timeline could be either after the camera-ready phase, or after the conference, or after the publication of the paper in the digital libraries. For instance, Social Science Research Network (SSRN) with minimal vetting allows updates to the paper in the post-publication phase. Certain fields (such as economics) have a notion of working papers instead. Examples of post-publication updates could include either replication concerns or corrections (or extensions) by authors themselves. Journals allow such mechanisms where the editor-in-chief mediates the communication with the authors and may eventually make the decision to publish the letter with (or without) authors' response. The model for conferences is unclear and a mechanism for vetting and mediation needs to be defined. The Association for Computing Machinery (ACM) has a mechanism to post comments on the DL. The Internet Engineering Task Force (IETF) has a mechanism to associate errata with immutable Request for Comments (RFC) since 2002 which are mediated by the IETF Area Director (AD). ACM SIGCOMM CCR also has a similar errata mechanism. A strawman proposal is for the conference steering committee to designate a Point of Contact (PoC) to handle post-publication concerns and also deal with possible misuse scenarios that are mediated with the authors with responses tagged in DL. While, a counter viewpoint is to avoid hierarchical control structures, but instead crowd source the problem by utilising systems that already exist. For instance, StackExchange, which uses reputation metrics to rank most useful answers and ResearchGate, a social network for scientists that other scientific communities have meanwhile adopted.

4.8 Data and Metadata Formats

Ken Calvert, Kimberly Claffy, Alberto Dainotti, Ralph Holz, Daniel Karrenberg, Robert Kisteleki, Joel Sommers, Brian Trammell and John Wroclawski

The group met with the goal to understand how data (and metadata) can be represented in a common understandable way to lower the barriers of collaboration and facilitate reproducibility of research. One aspect that is of particular significance is metadata which allows for search, categorisation and retrieval of the raw data. The metadata describes the context within which measurements were performed, the experimental parameters and overall interpretation of the data and how it was created, and its ownership and access rights. As such, it is essential to treat metadata as a first-class citizen because allowing it to be an afterthought often leads to its neglect. With this in mind, the goal is to drive creation of common tooling of interchange formats and APIs to facilitate the consumer of the data. BGPstream [32] is one such case study (of an API and a library) that provides common access to BGP data from multiple providers. The solution is specifically tailored to BGP datasets only and has seen increasing uptake in the research community. The group wondered whether it would be possible to learn from this effort and build infrastructures for other kinds of datasets with similar success? `traceroute` datasets are one such target type of datasets that are produced by multiple providers: CAIDA Archipelago, RIPE Atlas and Measurement Lab measurement infrastructures. The group plans to take this discussion forward by attempting to build a BGPstream-like architecture for `traceroute` data.

5 Conclusions and Next Steps

Participants with a mix of senior and junior researchers hailing from both academia and industry encouraged fruitful dialogue. A number of future research agendas were recognized. The group working on hyperpapers (§3.1 and §4.1) submitted the idea and call for actions as an editorial note [19] for the SIGCOMM CCR. The AEC plans to review the artifacts of 32 submitted papers with the help of 37 volunteers who evalute the artifacts and prepare public reviews. A report of the AEC activity with public badges will appear in 2019. Damien Saucez lead the activity of badging accepted papers that released artifacts for the CoNEXT 2018 conference. 14 (out of 32 accepted) papers were submitted for evaluation, and 12 papers were badged with the help of 20 volunteers. The badges are marked on the conference webpage[4] and also in the ACM DL. The group working on preparing students to think about reproducibility (§4.2) prepared an editorial note [12] on a beginners guide to reproducibility for experimental networking research which recently appeared in SIGCOMM CCR. The group working on preparing a reproducibility guidelines for reviewers (§4.3) produced a review form which informed the reviewers of the AEC. The group working on a reproducibility track (§4.6) worked with Anja Feldmann who made a proposal for such a reproducibility track at IMC 2018[5].

The organizing team also received valuable feedback. The participants felt productive in the groups and appreciated the continuation of the group activity in the mornings. Participants appreciated that the talks were limited to seven minutes to allow more time for group discussions, but also suggested to encourage presenters to present on topics that increased interactivity.

References

1 ACM Artifact Review and Badging. https://www.acm.org/publications/policies/artifact-review-badging [Online; last accessed 29-December-2018].
2 Computer and Network Systems (CNS): Core Programs. https://www.nsf.gov/pubs/2018/nsf18569/nsf18569.htm [Online; last accessed 29-December-2018].
3 Elsevier SoftwareX. https://www.journals.elsevier.com/softwarex [Online; last accessed 29-December-2018].
4 IMPACT Cyber Trust. https://www.impactcybertrust.org [Online; last accessed 31-December-2018].
5 NSF Proposal & Award Policies & Procedures Guide (PAPPG). https://www.nsf.gov/pubs/policydocs/pappg18 [Online; last accessed 29-December-2018].
6 TMA PhD School on Traffic Monitoring and Analysis. http://tma.ifip.org/2017/phd-school. [Online; accessed 25-December-2018].

4 https://conferences2.sigcomm.org/co-next/2018
5 https://conferences.sigcomm.org/imc/2018

7 Toward a Compatible Reproducibility Taxonomy for Computational and Computing Sciences. https://cfwebprod.sandia.gov/cfdocs/CompResearch/docs/SAND2018-11186.pdf [Online; last accessed 31-December-2018].

8 ACM Task Force on Data, Software and Reproducibility in Publication, 2015. https://www.acm.org/publications/task-force-on-data-software-and-reproducibility [Online; last accessed 29-December-2018].

9 Özgü Alay, Andra Lutu, Miguel Peón Quirós, Vincenzo Mancuso, Thomas Hirsch, Kristian Evensen, Audun Fosselie Hansen, Stefan Alfredsson, Jonas Karlsson, Anna Brunström, Ali Safari Khatouni, Marco Mellia, and Marco Ajmone Marsan. Experience: An Open Platform for Experimentation with Commercial Mobile Broadband Networks. In *MOBICOM*, pages 70–78, 2017. doi:10.1145/3117811.3117812.

10 Alberto Dainotti. A Conversation on Hyperpapers and Open Co-authoring. http://www.caida.org/workshops/aims/1803 [Online; last accessed 31-December-2018].

11 Vaibhav Bajpai, Arthur W. Berger, Philip Eardley, Jörg Ott, and Jürgen Schönwälder. Global Measurements: Practice and Experience (Report on Dagstuhl Seminar #16012). *Computer Communication Review*, 46(2):32–39, 2016. doi:10.1145/2935634.2935641.

12 Vaibhav Bajpai, Anna Brunström, Anja Feldmann, Wolfgang Kellerer, Aiko Pras, Henning Schulzrinne, Georgios Smaragdakis, Matthias Wählisch, and Klaus Wehrle. The dagstuhl beginners guide to reproducibility for experimental networking research. *SIGCOMM Computer Communication Review (CCR)*, 49(1):24–30, 2019. doi:10.1145/3314212.3314217.

13 Vaibhav Bajpai, Steffie Jacob Eravuchira, Jürgen Schönwälder, Robert Kisteleki, and Emile Aben. Vantage Point Selection for IPv6 Measurements: Benefits and Limitations of RIPE Atlas Tags. In *IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, pages 37–44, 2017. doi:10.23919/INM.2017.7987262.

14 Vaibhav Bajpai, Mirja Kühlewind, Jörg Ott, Jürgen Schönwälder, Anna Sperotto, and Brian Trammell. Challenges with Reproducibility. In *SIGCOMM 2017 Reproducibility Workshop*, pages 1–4, 2017. doi:10.1145/3097766.3097767.

15 Vaibhav Bajpai and Jürgen Schönwälder. A Survey on Internet Performance Measurement Platforms and Related Standardization Efforts. In *IEEE Communications Surveys and Tutorials*, pages 1313–1341, 2015. doi:10.1109/COMST.2015.2418435.

16 Adam Brinckman, Kyle Chard, Niall Gaffney, Mihael Hategan, Matthew B. Jones, Kacper Kowalik, Sivakumar Kulasekaran, Bertram Ludäscher, Bryce D. Mecum, Jarek Nabrzyski, Victoria Stodden, Ian J. Taylor, Matthew J. Turk, and Kandace Turner. Computing Environments for Reproducibility: Capturing the "Whole Tale". In *Future Generation Computer Systems*, 2018. doi:10.1016/j.future.2017.12.029.

17 Bryan Clark, Todd Deshane, Eli M. Dow, Stephen Evanchik, Matthew Finlayson, Jason Herne, and Jeanna Neefe Matthews. Xen and the Art of Repeated Research. In *USENIX Annual Technical Conference*, pages 135–144, 2004. URL: http://www.usenix.org/publications/library/proceedings/usenix04/tech/freenix/clark.html.

18 Christian S. Collberg and Todd A. Proebsting. Repeatability in Computer Systems Research. In *http://doi.acm.org/10.1145/2812803*, Communications of the ACM, 2016. doi:10.1145/2812803.

19 Alberto Dainotti, Ralph Holz, Mirja Kühlewind, Andra Lutu, Joel Sommers, and Brian Trammell. Open collaborative hyperpapers: a call to action. *SIGCOMM Computer Communication Review (CCR)*, 49(1):31–33, 2019. doi:10.1145/3314212.3314218.

20 Philip Eardley, Marco Mellia, Jörg Ott, Jürgen Schönwälder, and Henning Schulzrinne. Global Measurement Framework (Dagstuhl Seminar 13472). *Dagstuhl Reports*, 3(11):144–153, 2013. doi:10.4230/DagRep.3.11.144.

21 Paul Emmerich, Sebastian Gallenmüller, Daniel Raumer, Florian Wohlfart, and Georg Carle. MoonGen: A Scriptable High-Speed Packet Generator. In *Internet Measurement Conference (IMC)*, pages 275–287, 2015. doi:10.1145/2815675.2815692.

22 Matthias Flittner, Robert Bauer, Amr Rizk, Stefan Geissler, Thomas Zinner, and Martina Zitterbart. Taming the Complexity of Artifact Reproducibility. In *SIGCOMM 2017 Reproducibility Workshop*, 2017. doi:10.1145/3097766.3097770.

23 Matthias Flittner, Mohamed Naoufal Mahfoudi, Damien Saucez, Matthias Wählisch, Luigi Iannone, Vaibhav Bajpai, and Alex Afanasyev. A Survey on Artifacts from CoNEXT, ICN, IMC, and SIGCOMM Conferences in 2017. In *SIGCOMM Computer Communication Review*. ACM, April 2018. doi:10.1145/3211852.3211864.

24 Fabien Geyer and Georg Carle. Network Engineering for Real-time Networks: Comparison of Automotive and Aeronautic Industries Approaches. *IEEE Communications Magazine*, pages 106–112, 2016. doi:10.1109/MCOM.2016.7402269.

25 Thomas Holterbach, Emile Aben, Cristel Pelsser, Randy Bush, and Laurent Vanbever. Measurement Vantage Point Selection Using A Similarity Metric. In *Applied Networking Research Workshop*, pages 1–3, 2017. doi:10.1145/3106328.3106334.

26 Heidi Howard, Malte Schwarzkopf, Anil Madhavapeddy, and Jon Crowcroft. Raft Refloated: Do We Have Consensus? In *Operating Systems Review*, pages 12–21, 2015. doi:10.1145/2723872.2723876.

27 Catherine Jones, Brian Matthews, and Ian Gent. Software reuse, repurposing and reproducibility. In *Proceedings of the 12th International Conference on Digital Preservation, iPRES 2015, Chapel Hill, North Carolina, USA, November 2-6, 2015*, 2015. URL: http://hdl.handle.net/11353/10.429590.

28 Balachander Krishnamurthy, Walter Willinger, Phillipa Gill, and Martin F. Arlitt. A Socratic Method for Validation of Measurement-based Networking Research. In *Computer Communications*, pages 43–53, 2011. doi:10.1016/j.comcom.2010.09.014.

29 Stuart Kurkowski, Tracy Camp, and Michael Colagrosso. MANET Simulation Studies: The Incredibles. In *Mobile Computing and Communications Review*, pages 50–61, 2005. doi:10.1145/1096166.1096174.

30 Anna Maria Mandalari, Andra Lutu, Ana Custura, Ali Safari Khatouni, Özgü Alay, Marcelo Bagnulo, Vaibhav Bajpai, Anna Brunström, Jörg Ott, Marco Mellia, and Gorry Fairhurst. Experience: Implications of Roaming in Europe. In *MOBICOM*, pages 179–189, 2018. doi:10.1145/3241539.3241577.

31 National Academies of Sciences, Engineering, and Medicine. *Open Science by Design: Realizing a Vision for 21st Century Research*. The National Academies Press, 2018. doi:10.17226/25116.

32 Chiara Orsini, Alistair King, Danilo Giordano, Vasileios Giotsas, and Alberto Dainotti. BGPStream: A Software Framework for Live and Historical BGP Data Analysis. In *Internet Measurement Conference (IMC)*, 2016. URL: http://dl.acm.org/citation.cfm?id=2987482.

33 Vern Paxson. Strategies for Sound Internet Measurement. In *Internet Measurement Conference, (IMC)*, pages 263–271, 2004. doi:10.1145/1028788.1028824.

34 Diana Andreea Popescu and Andrew W. Moore. Reproducing Network Experiments in a Time-controlled Emulation Environment. In *Traffic Monitoring and Analysis (TMA)*, 2016. URL: http://tma.ifip.org/2016/papers/tma2016-final10.pdf.

35 Andreas Reuter, Randy Bush, Ítalo S. Cunha, Ethan Katz-Bassett, Thomas C. Schmidt, and Matthias Wählisch. Towards a Rigorous Methodology for Measuring Adoption of RPKI Route Validation and Filtering. In *SIGCOMM Computer Communication Review*, 2018. doi:10.1145/3211852.3211856.

36 Laura Rueda, Martin Fenner, and Patricia Cruse. Datacite: Lessons learned on persistent identifiers for research data. *IJDC*, 11(2):39–47, 2016. doi:10.2218/ijdc.v11i2.421.

37 Geir Kjetil Sandve, Anton Nekrutenko, James Taylor, and Eivind Hovig. Ten Simple Rules for Reproducible Computational Research. *PLOS Computational Biology Journal*, 2013. doi:10.1371/journal.pcbi.1003285.

38 Damien Saucez and Luigi Iannone. Thoughts and Recommendations from the ACM SIGCOMM 2017 Reproducibility Workshop. *SIGCOMM Computer Communication Review*, 2018. doi:10.1145/3211852.3211863.

39 Quirin Scheitle, Matthias Wählisch, Oliver Gasser, Thomas C. Schmidt, and Georg Carle. Towards an Ecosystem for Reproducible Research in Computer Networking. In *SIGCOMM Reproducibility Workshop*, pages 5–8, 2017. doi:10.1145/3097766.3097768.

40 Henning Schulzrinne. Networking Research - A Reflection in the Middle Years. In *Computer Communications*, 2018. doi:10.1016/j.comcom.2018.07.001.

41 Patrick Vandewalle, Jelena Kovacevic, and Martin Vetterli. Reproducible Research in Signal Processing. *IEEE Signal Processing Magazine*, 2009. doi:10.1109/MSP.2009.932122.

42 Lisa Yan and Nick McKeown. Learning Networking by Reproducing Research Results. In *Computer Communication Review*, 2017. doi:10.1145/3089262.3089266.

43 Jihwang Yeo, David Kotz, and Tristan Henderson. CRAWDAD: A Community Resource for Archiving Wireless Data at Dartmouth. *SIGCOMM Computer Communication Review*, pages 21–22, 2006. doi:10.1145/1129582.1129588.

Participants

- Vaibhav Bajpai
TU München, DE
- Steve Bauer
MIT – Cambridge, US
- Olivier Bonaventure
UC Louvain, BE
- Anna Brunström
Karlstad University, SE
- Kenneth L. Calvert
University of Kentucky – Lexington, US
- Georg Carle
TU München, DE
- Kimberly Claffy
San Diego Supercomputer Center, US
- Alberto Dainotti
San Diego Supercomputer Center, US
- Anja Feldmann
MPI für Informatik – Saarbrücken, DE
- Ralph Holz
The University of Sydney, AU
- Luigi Iannone
Télécom Paris Tech, FR
- Daniel Karrenberg
RIPE NCC – Amsterdam, NL
- Wolfgang Kellerer
TU München, DE
- Robert Kisteleki
RIPE NCC – Amsterdam, NL
- Mirja Kühlewind
ETH Zürich, CH
- Andra Lutu
Telefónica Research – Barcelona, ES
- Matt Mathis
Google Inc. – Mountain View, US
- Jörg Ott
TU München, DE
- Aiko Pras
University of Twente, NL
- Damien Saucez
INRIA Sophia Antipolis, FR
- Quirin Scheitle
TU München, DE
- Jürgen Schönwälder
Jacobs University Bremen, DE
- Henning Schulzrinne
Columbia University – New York, US
- Georgios Smaragdakis
TU Berlin, DE
- Karen Sollins
MIT – Cambridge, US
- Joel Sommers
Colgate University – Hamilton, US
- Brian Trammell
ETH Zürich, CH
- Matthias Wählisch
FU Berlin, DE
- Klaus Wehrle
RWTH Aachen University, DE
- John Wroclawski
USC – Marina del Rey, US
- Thomas Zinner
TU Berlin, DE