

AI for Social Good

Claudia Clopath^{*1}, Ruben De Winne^{*2},
Mohammad Emtiyaz Khan^{*3}, and Jacopo Margutti^{*4}

1 Imperial College London, GB. c.clopath@imperial.ac.uk

2 Oxfam Novib – The Hague, NL. ruben_de_winne@hotmail.com

3 RIKEN – Tokyo, JP. emtiyaz.khan@riken.jp

4 510 / Netherlands Red Cross – The Hague, NL. jmargutti@redcross.nl

Abstract

Progress in the field of Artificial intelligence (AI) and machine learning (ML) has not slowed down in recent 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. AI for Social Good in general is trying to ensure that the social good does not become an afterthought, but that society benefits as a whole. In this Dagstuhl Seminar, which can be considered a follow-up edition of Dagstuhl Seminars 19082 and 22091 with the same title, we brought together AI and machine learning researchers with non-governmental organisations (NGOs), as they already pursue a social good goal, have rich domain knowledge, and vast networks with (non-)governmental actors in developing countries. Such collaborations 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. In this seminar, researchers and practitioners from diverse areas of machine learning joined stakeholders from a range of NGOs to spend a week together. We first pursued an improved understanding of each side's challenges and established a common language, via presentations and discussion groups. Building on this foundation, we organised a hackathon around some existing technical questions within the NGOs to scope the applicability of AI methods and seed collaborations. Finally, we discussed topics that cut across the AI for social good field, such as how to properly evaluate AI models that are used for good.

Seminar February 18–23, 2024 – <https://www.dagstuhl.de/24082>


2012 ACM Subject Classification Computing methodologies → Artificial intelligence; Applied computing → Computers in other domains; Computing methodologies → Machine learning

Keywords and phrases artificial intelligence, interdisciplinary, machine learning, non-governmental organizations, social good

Digital Object Identifier 10.4230/DagRep.14.2.182

1 Executive Summary

Ruben De Winne (Oxfam Novib – The Hague, NL)

License  Creative Commons BY 4.0 International license
© Ruben De Winne

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, and more recently via general uptake of software

* Editor / Organizer



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 4.0 International license

AI for Social Good, *Dagstuhl Reports*, Vol. 14, Issue 2, pp. 182–190

Editors: Claudia Clopath, Ruben De Winne, Mohammad Emtiyaz Khan, and Jacopo Margutti



DAGSTUHL
REPORTS

Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

applications built on large language models. 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 pursuing various social good goals, such as providing legal aid, providing humanitarian assistance, advocating for gender justice, denouncing growing levels of inequality, 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 & 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 & discussions on the one hand with a hands-on hackathon on the other hand. High-level talks & 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 generalizability, data pipelines, and explainability. These talks and discussions allowed all participants – in a very short timeframe – 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 to increase 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 topics that cut across the AI for social good field, such as how to properly evaluate AI models that are used for good.

2 Table of Contents

Executive Summary
Ruben De Winne 182

Overview of Talks

Potential Generic Tools for AI for Social Good: Multi-objective Optimization & AutoML
Frank Hutter 185

Method seeds
Mohammad Emtiyaz Khan 185

AI for Data, Models, Decisions
Subhransu Maji 186

A brief history of NLP
Virginia Patridge 186

Machine Learning for Peace: Digital Tools for Civic Actors
Jeremy Springman 186

Working groups

Sustainability and ownership of AI models
Ruben De Winne 187

Evaluation of AI models
Daphne Ezer and Ruben De Winne 188

Panel discussions

Feedback for the organizers of a follow-up AI for Social Good seminar at Dagstuhl
Claudia Clopath 189

Participants 190

3 Overview of Talks

3.1 Potential Generic Tools for AI for Social Good: Multi-objective Optimization & AutoML


Frank Hutter (Universität Freiburg, DE)

License  Creative Commons BY 4.0 International license
© Frank Hutter

Professor Hutter presented the potential of AutoML for Multi-objective Optimization in the context of AI for Social Good. Setting the scene, professor Hutter used the example problem of algorithmic fairness in face recognition. Facial recognition systems are widely acknowledged to be prone to bias, particularly along sociodemographic dimensions such as gender and race. Given their pervasive use in sensitive applications like law enforcement for suspect identification and missing person tracking, there is a pressing need to address this bias. Professor Hutter explored the challenge of improving the fairness of face recognition algorithms while maintaining high accuracy rates. A key strategy discussed is multi-objective optimization, which involves balancing competing objectives such as minimizing errors while reducing bias. This approach mirrors similar tradeoffs found in other domains like food production and language generation algorithms such as GPT. By leveraging multi-objective AutoML (Automated Machine Learning), it becomes possible to develop AI systems that are not only performant but also fair, calibrated, energy-efficient, and robust. The advantages of AutoML are highlighted, including its ability to streamline ML application development, ensure reproducibility and transparency, and potentially surpass human performance on various tasks. Also, especially relevant in the AI for Social Good context, is the potential of unblocking applications because AutoML can remove the requirement for a (scarce, highly-paid) human ML expert in the inner loop. However, it's cautioned that while tools for single-objective tasks and tabular data are mature, multi-objective AutoML tooling is still somewhat in its early stages. Moreover, the presentation emphasized the importance of considering the broader socio-technical context in which AI systems operate. While technical solutions like AutoML offer promise, they must be integrated thoughtfully within the larger socio-technical system to truly address issues of fairness and bias in face recognition and other AI applications.

3.2 Method seeds


Mohammad Emtiyaz Khan (RIKEN – Tokyo, JP)

License  Creative Commons BY 4.0 International license
© Mohammad Emtiyaz Khan

Mainly for the NGO participants who lacked an in-depth comprehension of what Machine Learning entails, professor Khan gave a talk on the main ML concepts and methods. He suggested his definition of ML (“making a computer intelligent without explicitly programming it”), noted a number of historical and recent success stories as well as notorious failure cases, and made participants aware of some key challenges (e.g. existing methods require a large amount of ‘good-quality’ data). He also explained main divisions within the ML realm and the most important methods ((un)supervised learning, reinforcement learning, ...).

3.3 AI for Data, Models, Decisions

Subhransu Maji (University of Massachusetts – Amherst, US)

License  Creative Commons BY 4.0 International license
© Subhransu Maji

Professor Maji started his presentation by explaining that billions of birds migrate every year, mostly in the cover of darkness. But these birds are visible on RARD networks in the continental US. Subhransu explained that using Machine Learning, we can learn how migration has changed over the last 25+ years. A team of ecologists and computer scientists worked together to analyze this bird migration data at scale. Challenges and unique opportunities that this collaboration had were also discussed. Subhransu concluded with three pieces of advice to the participants:

- Don't throw away noisy data! You might be able to correct for noise.
- Don't throw away info on who labeled data!
- Don't throw away intermediate things, might be useful at some point for training.

Professor Maji also highlighted the iNaturalist application of AI for biodiversity mapping. He asked the question what if AI is not reliable? Many applications require a total count – e.g., how many birds migrate in a year, or how many damaged buildings are in city / county / state. Estimates can be biased (sometimes by a lot!).

3.4 A brief history of NLP

Virginia Partridge (University of Massachusetts – Amherst, US)

License  Creative Commons BY 4.0 International license
© Virginia Partridge

To complement professor Khan's presentation, Virginia Partridge zoomed in on the topic of Natural Language Processing (NLP). She explained the difference between generative and predictive NLP tasks, gave an overview of the history of NLP, and walked participants through an overview of NLP models.

3.5 Machine Learning for Peace: Digital Tools for Civic Actors

Jeremy Springman (University of Pennsylvania – Philadelphia, US)

License  Creative Commons BY 4.0 International license
© Jeremy Springman

Dr. Springman presented Digital Tools for Civic Actors from UPenn's Machine Learning for Peace lab. In general, the lab's approach for using data to contribute to crisis response is the following:


1. Awareness: data on what's happening very recently
 - Mass scraping online news + ML to track events
 - Interactive data dashboard
2. Planning: predictive analytics for strategic decisions
 - Forecasting political events
 - Civic Space Early Warning System

The lab uses online news from 300+ sources in 35 languages as input data. To ensure data quality, the focus is on reputable local sources. Overall, there is much better coverage than extant archivers/aggregators (GDELT, Wayback, Lexis Nexis, etc.). Dr. Springman concluded with a few concrete examples of detecting and forecasting civic events, such as arrests in Uganda or civic activism in Angola. In the future, the lab wants to be capable of forecasting external event data (Travel Advisories) and of extracting new information from text.

4 Working groups

4.1 Sustainability and ownership of AI models

Ruben De Winne (Oxfam Novib – The Hague, NL)

License  Creative Commons BY 4.0 International license
© Ruben De Winne

1. Common failures: A typical failure scenario involves enthusiasm during deployment, followed by eventual abandonment due to unforeseen issues and lack of ongoing usage.
2. Knowledge loss: High turnover rates, typically every 2-3 years, contribute to knowledge loss within organizations, affecting the sustainability of AI projects.
3. Maintenance challenges: While building prototypes may be exciting for students/researchers, finding skilled software engineers for basic maintenance is difficult, highlighting the importance of technical expertise.
4. NGO ownership: NGOs possessing the right profiles and resources can successfully own and maintain AI projects.
5. Business models: Offering digital products as services to other organizations can be a sustainable business model for AI applications.
6. Avoiding dependencies: Careful selection of services is crucial to avoid dependencies on technical partners who may package solutions in a way that limits maintenance without their continued support.
7. Partnerships for trust and resources: Partnerships with academia and other organizations can enhance trust, provide resources, and facilitate peer review of methods.
8. Communities of practice: Engaging in communities such as NetHope can provide opportunities for brainstorming and accessing resources.
9. University partnerships: Collaborating with universities, as seen with KoboToolbox, can yield successful outcomes in AI project sustainability.
10. Focus on reusable tools: Reusable tools are easier to manage and sustain as their impact can be easily demonstrated, facilitating resource allocation and support.

4.2 Evaluation of AI models

Daphne Ezer (University of York, GB) and Ruben De Winne (Oxfam Novib – The Hague, NL)

License  Creative Commons BY 4.0 International license
© Daphne Ezer and Ruben De Winne

In the session on evaluating AI models, several key points were discussed:

- **NGO Concerns:** NGOs are concerned about ensuring AI does no harm.
- **Metrics for Chatbots:** There's ambiguity regarding evaluation metrics for Generative AI technologies like chatbots. Questions arise about whether chatbots are effectively answering questions and how well they're performing overall.
- **Building Benchmarks:** Developing benchmarks for chatbots involves defining basic proxy metrics such as factuality, toxicity, and relevance. However, the ultimate measure of success lies in user actions, like registration.
- **A/B Testing:** A/B testing involves comparing different versions of the chatbot to measure user engagement, typically assessed by how long users interact with it.
- **Multi-Objective Optimization:** To monitor progress towards goals, it's essential to define and measure multiple metrics specific to the use case.
- **Avoiding Bias:** Defining metrics before designing methods or models helps avoid confirmation bias.
- **Iterative Improvement:** Expect multiple iterations of testing and improvement to refine AI models.
- **Beware of Overfitting:** Using publicly available benchmarks might lead to overestimating model performance if the model has been trained on those benchmarks.
- **Monitoring Performance:** Implement alerts to detect significant changes in prediction distributions, especially for categorical outputs.
- **Readiness for Production:** Models are deemed ready for production based on predefined metrics, followed by extensive qualitative checks.
- **Domain Knowledge:** Evaluation now requires more domain knowledge than traditional feature engineering.
- **Categorizing Errors:** Errors should be categorized based on their impact, distinguishing between manageable and catastrophic mistakes.
- **Focus on Worst Outcomes:** Pay attention to extreme cases rather than just averages when evaluating model performance.
- **Task-Specific Evaluation:** Evaluation frameworks need to be tailored to specific tasks, making generalizable frameworks challenging to create.

5 Panel discussions

5.1 Feedback for the organizers of a follow-up AI for Social Good seminar at Dagstuhl

Claudia Clopath (Imperial College London, GB)

License  Creative Commons BY 4.0 International license
© Claudia Clopath

- Don't change too much because it works really good.
- Keep the pre-call between AI participants and NGO participants.
- Ensure that everyone is at the same starting point, e.g. by being clear on problem, sharing reading material on what ML is or Nature Comms paper, have a checklist for the prep call (e.g. re: data, problem, ...).
- Develop an MOOC or similar to make sure that everyone is at the same level.
- Be clear beforehand what the expectations are, so that people can prepare properly (explicitly that actually bringing data to the seminar could be helpful).
- Good to have a general introduction (Data science lifecycle, demystifying AI) session to start the week.
- Develop glossary in advance, maybe turn it into a fun quiz in the beginning.
- The case studies may be presented with some slides instead that only in an oral form.

Participants

- Asma Atamna
Ruhr-Universität Bochum, DE
- Annabelle Behnke
Deutsches Rotes Kreuz e.V. –
Berlin, DE
- Siu Lun Chau
CISPA – Saarbrücken, DE
- Claudia Clopath
Imperial College London, GB
- Jorn Dallinga
WWF – Zeist, NL
- Ruben De Winne
Oxfam Novib – The Hague, NL
- Michael Dhatemwa
Oxfam Novib – The Hague, NL
- Daphne Ezer
University of York, GB
- Elshad Gojayev
Internews – Washington, US
- Frank Hutter
Universität Freiburg, DE
- Roberto Interdonato
CIRAD – Montpellier, FR
- Mohammad Emtiyaz Khan
RIKEN – Tokyo, JP
- Isabell Klipper
Deutsches Rotes Kreuz e.V. –
Berlin, DE
- Parvathy Krishnan
Analytics for a Better World –
Amsterdam, NL
- Derek Loots
Médecins Sans Frontières –
Amsterdam, NL
- Subhransu Maji
University of Massachusetts –
Amherst, US
- Jacopo Margutti
510 / Netherlands Red Cross –
The Hague, NL
- Marieke Meeske
Tilburg University, NL & Oxfam
Novib – Den Haag, NL
- Krikamol Muandet
CISPA – Saarbrücken, DE
- Virginia Partridge
University of Massachusetts –
Amherst, US
- Julia Proskurnia
Google – Zürich, CH
- Lennart Purucker
Universität Freiburg, DE
- Jake Robertson
Universität Freiburg, DE
- Andrés Roure Cuzzoni
Propel – Den Haag, NL
- Tom Schaul
Google DeepMind – London, GB
- Jeremy Springman
University of Pennsylvania –
Philadelphia, US
- Maïna Vergonjanne
Droits Quotidiens Legal Tech –
Montpellier, FR

