Data-driven Humanitarian Mapping and Policymaking: Toward Planetary-Scale Resilience, Equity, and Sustainability

Data-driven Humanitarian Mapping and Policymaking: Toward Planetary-Scale Resilience, Equity, and Sustainability

Snehalkumar 'Neil' Gaikwad*
Massachusetts Institute of Technology
Cambridge, USA

Shankar Iyer
Meta Research
Menlo Park, USA

Dalton Lunga
Oak Ridge National Laboratory
Oak Ridge, USA

Takahiro Yabe
Massachusetts Institute of Technology
Cambridge, USA

Xiaofan Liang
Georgia Institute of Technology
Atlanta, USA

Bhavani Ananthabhotla
Massachusetts Institute of Technology
Cambridge, USA

Nikhil Behari
Harvard University
Cambridge, USA

Sreelekha Guggilam
Oak Ridge National Laboratory
Oak Ridge, USA

Guanghua Chi
Meta Research
Menlo Park, USA

MOTIVATION

Human civilization faces existential threats in the forms of climate change, food insecurity, pandemics, international conflicts, forced displacements, and environmental injustice. These overarching humanitarian challenges disproportionately impact historically marginalized communities worldwide. UN OCHA estimates that 274 million people will need humanitarian support in 2022.[1] Despite growing perils to human and environmental well-being, there remains a paucity of publicly-engaged computing research to inform the design of interventions. Data science and machine learning efforts exist, but they remain isolated from socioeconomic, environmental, cultural, and policy contexts at local and international scales. Moreover, biases and privacy infringements in data-intensive methods amplify existing inequalities and harms. The result is that proclaimed benefits of data-driven innovations may remain inaccessible to policymakers, practitioners, and underserved communities whose lives they intend to transform. To address gaps in knowledge and improve the livelihood of marginalized populations, we have established the Data-driven Humanitarian Mapping and Policymaking, an interdisciplinary initiative [1–3].

VISION: EQUITABLE DATA SCIENCE AND POLICYMAKING FOR GLOBAL INCLUSION

We envision ethical designs of novel data science methods, community partnerships, and evidence-based policies for global-scale resilience, equity, and sustainability. Over the last few decades, the world has seen unprecedented growth in data-intensive systems and methods (such as federated machine learning, graph neural networks, privacy-preserving computations, transfer learning, earth observation remote sensing, mobile phone data, online social media, surveys, crowdsourcing datasets, etc.). Socially and ethically responsible designs of these analytics tools hold the potential to advance our understanding of the spatiotemporal complexities and risks underpinning humanitarian and sustainability policy challenges. Data-driven Humanitarian Mapping and Policymaking drives the design, engineering, and deployment of humanitarian technologies through computational, legal, ethical, and policy lenses [1–3]. Specifically, we aim to design, develop, and critically evaluate: (1) **data ecosystems:** novel and ethical ways to collect, share, and validate multimodal humanitarian, environmental, and socioeconomic, development data; (2) **data-driven methodologies and systems for decision-making:** identify, map, and measure climate crisis, human displacements, pandemics, human trafficking, food insecurity, socioeconomic development, poverty, racial and gender-based violence, natural disasters, biodiversity conservation, at-risk communities, human rights violations, infrastructure damages, malaria prevalence, and environmental injustice, as a case in point; (3) **trustworthy designs of AI and community-based governance mechanisms:** mitigating algorithmic harms by incorporating privacy, fairness, interpretability, accountability, transparency, and ethics; and (4) **action research and accountable dissemination of insights:** evidence-based actionable insights for humanitarian responders, global development networks, NGOs, and policymakers. Each of these goals remains instrumental in fostering a global community of diverse stakeholders to advance a shared data science research agenda for inclusive humanitarian action.

KEYWORDS

data-intensive computing and society, algorithmic decision-making, AI ethics,data science and public policy, community-based design, human-AI collaboration, climate change, computational sustainability, public interest technology, remote sensing, data systems.

ACM Reference Format:
Snehalkumar 'Neil' Gaikwad, Shankar Iyer, Dalton Lunga, Takahiro Yabe, Xiaofan Liang, Bhavani Ananthabhotla, Nikhil Behari, Sreelekha Guggilam, and Guanghua Chi. 2022. Data-driven Humanitarian Mapping and Policymaking: Toward Planetary-Scale Resilience, Equity, and Sustainability. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '22), August 14–18, 2022, Washington, DC, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3534678.3542918

*Lead & corresponding author, gaikwad@mit.edu
[1] *UN OCHA statistics, accessed June 09, 2022*

KDD '22, August 14–18, 2022, Washington, DC, USA

ACM ISBN 978-1-4503-9385-0/22/08.
https://doi.org/10.1145/3534678.3542918

KDD WORKSHOP SERIES

As part of the initiative, we have hosted an annual workshop series at the SIGKDD conference since 2020 [1,2,3]. [2] The series provides a global platform to (1) share methodological advances in data science and policy to address pressing humanitarian challenges; (2) conceptualize design frameworks for data systems, federated ML, transfer learning, remote sensing, social media analysis, and edge computing for equitable high-stake decisions; (3) foster effective cross-disciplinary alliances and address risks of data-driven interventions in humanitarian actions and sustainable development.

The inception of the workshop series sparked from conversations between Neil Gaikwad, Shankar Iyer, Alex Dow, Carlos Castillo in 2019. The initiative is led by Neil Gaikwad with support from the organizing, advisory, and program committees. To date, the workshops were composed of a interdisciplinary group of 105 program committee members, 15 steering committee members, 12 keynote speakers and panelists, 50 contributions in the form of research papers and data systems tools, and over 200 participants worldwide. The keynote talks and panel discussions focused on data-driven humanitarian mapping in poverty alleviation, climate crisis, armed conflicts, food security, disaster response, equity, and the COVID-19 pandemic. Unifying themes across keynotes, talks, and panel discussions highlighted a range of research challenges that currently stand in the way of effectively translating data from measurement to privacy-preserving mapping, to federated machine learning, to model interpretability, and finally, to policy engagement.

3rd KDD Workshop on Data-driven Humanitarian Mapping and Policymaking. As a part of the third workshop, we will build on the prior success and continue expanding the vibrant community. We emphasize the theme of comprehensive global mapping of Earth Systems and Human Societies to address issues of bias, lack of representation, and data scarcity, especially for enabling environmental, social, and humanitarian justice. Humanitarian mapping initiatives aim to map a quantity of interest as comprehensively as possible worldwide. Examples include the Malaria Atlas Project, Meta's High-Resolution Settlement Layer, and various measures of the progression of the COVID-19 pandemic. The pandemic distinctly reminds us that we must coordinate at regional and global scales to respond to sustainability perils effectively. We will convene researchers and humanitarians to discuss the specific challenges while striving for full geographic coverage. How should practitioners navigate tradeoffs between spatial resolution and representativeness? How can practitioners ensure fair representations in global datasets? How can such datasets be validated? What are the risks associated with data-driven decision-making? We will discuss these and other related questions focused on human-machine autonomy through interdisciplinary perspectives.

We remain committed to promoting the values of diversity, equity, inclusion, and belonging in data science, computing, and policymaking. To increase participation at the workshop by underrepresented and historically marginalized groups, we have established a diversity scholarship. Additionally, the workshop guidelines require all participants to adhere to the anti-harassment policy.

[2] KDD Workshop on Data-driven Humanitarian Mapping and Policymaking

COMMUNITY

Speakers and Panelists. Sandy Pentland (MIT), Amy Rose (ORNL), Andrew Schroeder (Direct Relief), Caroline Buckee (Harvard), Catherine Nakalembe (University of Maryland), Dyan Mazurana (Tufts), Joshua Blumenstock (UC Berkeley), Justin Ginnetti (IDMC), Karen Chapple (UC Berkeley), Megan Price (HRDAG), Meha Jain (University of Michigan), Paige Maas (Meta), Sarah Williams (MIT)

Advisory Committee. Sandy Pentland (MIT), Alex Dow (Meta), Budhu Bhaduri (ORNL), Carlos Castillo (Univ. Pompeu Fabra), Haishan Fu (World Bank), Hamed Alemohammad (Radiant Earth), Jie Yin (University of Sydney), Joshua Blumenstock (UC Berkeley), Marta Gonzalez (UC Berkeley), Megan Price (HRDAG), Miho Mazereeuw (MIT), Rahul Panicker (Wadhwani AI), Ramesh Raskar (MIT), Vipin Kumar (University of Minnesota), Yu-Ru lin (University of Pittsburgh)

Program Committee. Abigail Horn, Aditya Mate, Akansha Saklani, Alex Dow, Alex Pompe, Amanda Coston, Amanda Hughes, Amanda Kube, Amy Rose, Andrés Abeliuk, Antonio Fernández Anta, Aparna Phalke, Aruna Sankaranarayanan, Ashique Khudabukhsh, Ayesha Mahmud, Bahman Rostami-Tabar, Belen Saldias Fuentes, Bianca Zadrozny, Björn Lütjens, Bryan Wilder, Carolynne Hultquist, Chintan Vaishnav, Ciro Cattuto, Clio Andris, Daniela Paolotti, David Reed, Delia Wendel, Emily Aiken, Eric Shook, Eric Sodomka, Esteban Moro, Eugenia Giraudy, Gautum Thakur, Girmaw Abebe Tadesse, Grigorios Kalliatakis, Hang Jiang, Hannah Kerner, Hao Li, Helene Benveniste, Hemant Purohit, Ingmar Weber, Ingmar Weber, Jackson Killian, Jigar Doshi, Jill Kelly, Joseph Aylett-Bullock, Karla Saldana Ochoa, Katherine Picchione, Katherine Hoffmann Pham, Kelly Jill, Kentaro Torisawa, Kevin Sparks, Kiran Garimella, Kostas Pelechrinis, Kristen Altenburger, Kuldeep Kurte, Kurt Luther, Laura Alessandretti, Lelia Hampton, Lijun Sun, Lily Xu, Marie Urban, Marie-Laure Charpignon, Marina Kogan, Marzia Rango, Meeyoung (Mia) Cha, Meha Jain, Mert Ertugrul, Minoo Rathnasabapathy, Morgan Frank, Muhammad Imran, Nagendra Singh, Nicera Wanjiru, Niklas Stoehr, Nishant Kishore, Nitin Kohli, Omari Sefu, Paige Maas, Philipe Dias, Priyank Lathwal, Rishemjit Kaur, Ritwik Gupta, Robert Soden, Sai Ravela, Sam Fraiberger, Samira Barzin, Satish Ukkusuri, Sefala Raesetje, Shawn Newsam, Sherrie Wang, Shruti Palaskar, Silvia Danielak, Siqi Zheng, Soroush Vosoughi, Tao Liu, Thilanka Munasinghe, Trivik Verma, Vanessa Frias-Martinez, Vedran Sekara, Vivek Singh, Wilhelmina Nekoto, Xidao Wen, Yan Leng, Yennie Jun, Yingjie Hu

REFERENCES

[1] S Gaikwad, S Iyer, Y Lin, and D Lunga. 2020. Proceedings of KDD 2021 Workshop on Data-driven Humanitarian Mapping: Harnessing Human-Machine Intelligence for High-Stake Public Policy and Resilience Planning. *arXiv preprint arXiv:2109.00435* (2020).

[2] S Gaikwad, S Iyer, D Lunga, and E Bondi. 2021. Data-driven Humanitarian Mapping: Harnessing Human-Machine Intelligence for High-Stake Public Policy and Resilience Planning. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. 4125–4126.

[3] S Gaikwad, S Iyer, D Lunga, and E Bondi. 2021. Proceedings of KDD 2021 Workshop on Data-driven Humanitarian Mapping: Harnessing Human-Machine Intelligence for High-Stake Public Policy and Resilience Planning. *arXiv preprint arXiv:2109.00100* (2021).