

# Data Management to Social Science and Back in the Future of Work

Sihem Amer-Yahia

sihem.amer-yahia@univ-grenoble-alpes.fr  
CNRS, Univ of Grenoble Alpes  
Grenoble, France

## Abstract

How will we work, live, and thrive in the post-pandemic future? The rapid mushrooming of online job markets has been transforming the definition of work and workplaces. After the pandemic, as we “cope with the new normal”, the future world of work may change forever and become predominantly virtual. This makes an unprecedented pool of talent available at our beck and calls to work on “gigs” that disband when the job is over; this also is the time of destabilization and changing nature of job security. As scientists, we have a big responsibility and a tremendous opportunity in shaping the Future of Work (FoW) post pandemic, by designing effective platforms that support productive employment, mitigate social costs, and provide an effective and safe learning environment.

A research agenda for FoW must mobilize the participation of various scientific, regulatory and miscellaneous stakeholders [10]. We will ask the questions: what is the role of Data Management (DM) in shaping research on FoW? Is now a ripe time to get Economics, Labor Theory, Psychology of Work and AI to help put DM research and technology at the center of research on FoW? Are we at all interested? The panelists will debate two complementary views: *A pessimistic view on whether FoW will tend to see humans as machines, robots, or low-level agents and use them in the service of broader AI goals vs. a more optimistic view, where AI and Social Science will help DM to develop technologies that empower humans for future workforce and workplaces.*

## CCS Concepts

• Information systems → Crowdsourcing.

## Keywords

Future of Work, Human Factors, Peer Learning, Gig Work, Working From Home

## ACM Reference Format:

Sihem Amer-Yahia and Senjuti Basu Roy. 2021. Data Management to Social Science and Back in the Future of Work. In *Proceedings of the 2021 International Conference on Management of Data (SIGMOD '21), June 20–25, 2021, Virtual Event, China*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3448016.3457536>

---

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

*SIGMOD '21, June 20–25, 2021, Virtual Event, China*

© 2021 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-8343-1/21/06.

<https://doi.org/10.1145/3448016.3457536>

Senjuti Basu Roy

senjutib@njit.edu  
New Jersey Institute of Technology  
Newark, NJ, USA

## 1 Significance

This proposal is motivated by a 3-day Shonan Meeting organized by the moderators in September 2019 that brought together researchers in human-in-the-loop Big Data, AI, and practitioners to connect the latest research findings to real-world questions and examine their impact on FoW. Follow up publications:

- i. **Shonan Report:** “Human-in-the-loop Big Data and AI: Connecting Theories and Practices for a Better Future of Work” [1].
- ii. **SIGMOD blog:** “Imagine All the People and AI in the FoW” [2].
- iii. **IEEE DEB Special Issue (Dec 2019):** “Imagine All the People and AI in the Future of Work” [3].
- iv. **SIGMOD record article:** “Data Management Research and More: Making AI Machines Work for Humans in FoW” [6].
- v. **IEEE DEB article published (Sep 2020):** “An ML-Powered Human Behavior Management System” [4].

## 2 Topic of Debate

The panelists will discuss the relevance of DM research to FoW. They will comment on the need for DM backends and for declarative languages to store and query work and workers, as well as the rising computational challenges, and how AI/ML will help design and optimize those backends.

**1. How might Social Science help DM capture human capabilities and social processes?** We envision that some of the critical challenges will lie in understanding different human roles in future jobs that are heavily AI-driven, modeling the inherent uncertainty in human behavior and adapting to their changing perceptions and needs. The development of interaction methodologies to provide onboarding and upskilling for workers, as well as enabling social interaction among workers would be of significant interest to the community.

**2. How might AI enable capturing and adapting to different stakeholder requirements?** The ability for all stakeholders in the new ecosystem, including regulators for policies and AI workers, to specify their requirements is another major challenge that must be achieved with declarative and high-level tools to express needs and expectations.

**3. How can we design appropriate DM ecosystems, benchmarks and metrics for FoW?** FoW will require overhauling the design and engineering of job platforms for collection, storage, retrieval, and analysis of the data deluge about workers, jobs, and their activities. This needs adaptive approaches to evolve the design of platforms and enhance computation capabilities. Benchmarking and development of appropriate metrics to measure system performance as well as human aspects (e.g., satisfaction, capital advancement, equity) is an additional challenge.

---

<sup>1</sup><https://wp.sigmod.org/?p=293>

**4. How can DM enforce multi-stakeholder fairness?** Regulations and policies should be developed to ensure the possibility of workers to unite and demand rights such as choosing among compensation mechanisms, and getting paid vacation. Similarly, platforms and job providers need to be given the means to safeguard themselves from abuse. Granting those rights implies the detection and reduction of discrimination in algorithmic decisions, especially in job allocation and exposure, through auditing processes and compliance to regulations [2].

## Confirmed Panelists

Our (confirmed) panelists will be invited to defend or debunk established and upcoming theories on FoW and on which research community will be best known for addressing FoW challenges:

- **Lei Chen** is a Chair Professor and the Director of HKUST Big Data Institute. Lei is a world expert in crowdsourcing-based data processing.
- **Krishna Gummadi** is a scientific director at the Max Planck Institute for Software Systems (MPI-SWS) in Germany. Krishna's research interests are in the measurement, analysis, design, and evaluation of complex Internet-scale systems. His work focus is on enhancing fairness and transparency of data-driven decision making in social computing systems.
- **Saiph Savage** is Director HCI Lab, WVU. She studies AI systems to help workers develop their digital skills to access better jobs and fight disinformation. Saiph was named one of the 35 Innovators under 35 by the MIT Technology Review for her civic tech research.
- **Jaime Teevan** is Chief Scientist for Microsoft's Experiences and Devices, and this past year coordinated the company's research efforts to understand how people's work practices have changed since the start of the pandemic: <http://aka.ms/newfutureofwork>.
- **Koichiro Yoshida** is the CEO of CrowdWorks (the largest crowdsourcing company in Japan, went to public in 2014). He will bring his expertise as a practitioner and a proponent of workers' rights in online marketplaces.

## Moderators

- **Sihem Amer-Yahia, CNRS, Univ. Grenoble Alpes, FRANCE** is a CNRS Research Director in Grenoble and the scientific deputy director of the Grenoble Computer Science Lab. Her interests are exploratory data analysis and fairness in job marketplaces. Before joining CNRS in 2012, she was Principal Scientist at QCRI, Senior Scientist at Yahoo! Research and Member of Technical Staff at at&t Labs. Sihem was Editor-in-Chief of the VLDB Journal, is co-chairing HILDA 2020, SIGMOD Demos 2020, ICDE 2021 and EDBT Demos 2021.
- **Senjuti Basu Roy, NJIT, USA** is an Assistant Professor at the New Jersey Institute of Technology. Her broader research interests lie in the area of data and content management with the focus on designing principled algorithms for “human-in-the-loop” systems. Senjuti co-chairs an IEEE international workshop on Human-in-the-loop Methods and Human Machine Collaboration in Big Data and has organized an NSF

workshop on converging human and technological perspectives in crowd-sourcing research.

- Sihem and Senjuti published numerous papers on online job marketplaces [7–9, 11–16], presented a tutorial at VLDB on the topic [5], and co-organized the Shonan Workshop on FoW in September 2019.

## References

- [1] 2019. Human-in-the-loop Big Data and AI: Connecting Theories and Practices for a Better Future of Work. In *Shonan Report, Meeting no 154*.
- [2] Heman Abdollahpouri and Robin Burke. 2019. Multi-stakeholder Recommendation and its Connection to Multi-sided Fairness. In *Proceedings of the Workshop on Recommendation in Multi-stakeholder Environments co-located with the 13th ACM Conference on Recommender Systems (RecSys 2019), Copenhagen, Denmark, September 20, 2019*.
- [3] Sihem Amer-Yahia, Lei Chen, Atsuyuki Morishima, and Senjuti Basu Roy. 2019. Special Issue on Imagine All the People and AI in the Future of Work. *IEEE Data Engineering Bulletin*.
- [4] Sihem Amer-Yahia, Reynold Cheng, Mohamed Bouadi, Abdelouahab Chibah, Mohammadreza Esfandiari, Jiangping Zhou, Nan Zhang, Eric Lau, Yuguo Li, Xiaolin Han, and Shivansh Mittal. 2020. An ML-Powered Human Behavior Management System. *43*, 3 (2020), 37–52.
- [5] Sihem Amer-Yahia and Senjuti Basu Roy. 2019. The Ever Evolving Online Labor Market: Overview, Challenges and Opportunities. *Proc. VLDB Endow.* *12*, 12 (2019), 1978–1981.
- [6] Sihem Amer-Yahia, Senjuti Basu Roy, Lei Chen, Atsuyuki Morishima, James Abello Monedero, Pierre Bourhis, François Charoy, Marina Danilevsky, Gautam Das, Gianluca Demartini, Shady Elbassuoni, David Gross-Amblard, Emilie Hoareau, Munenari Inoguchi, Jared B. Kenworthy, Itaru Kitahara, Dongwon Lee, Yunyao Li, Ria Mae Borromeo, Paolo Papotti, Raghav Rao, Sudeepa Roy, Pierre Senellart, Keishi Tajima, Saravanan Thirumuruganathan, Marion Tommasi, Kazutoshi Umemoto, Andrea Wiggins, and Koichiro Yoshida. 2020. Making AI Machines Work for Humans in FoW. *SIGMOD Rec.* *49*, 2 (2020), 30–35.
- [7] Sihem Amer-Yahia, Senjuti Basu Roy, Gautam Das, Joanna Lykourountzou, Habibur Rahman, and Saravanan Thirumuruganathan. 2018. Human Factors Modeling in Crowdsourcing. In *Encyclopedia of Database Systems, Second Edition*.
- [8] Mohammadreza Esfandiari, Senjuti Basu Roy, and Sihem Amer-Yahia. 2018. Explicit Preference Elicitation for Task Completion Time. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management, CIKM 2018, Torino, Italy, October 22–26, 2018*. 1233–1242.
- [9] Mohammadreza Esfandiari, Dong Wei, Sihem Amer-Yahia, and Senjuti Basu Roy. 2019. Optimizing Peer Learning in Online Groups with Affinities. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2019, Anchorage, AK, USA, August 4–8, 2019*. 1216–1226.
- [10] Aniket Kittur, Jeffrey V. Nickerson, Michael S. Bernstein, Elizabeth Gerber, Aaron D. Shaw, John Zimmerman, Matt Lease, and John J. Horton. 2013. The future of crowd work. In *Computer Supported Cooperative Work, CSCW 2013, San Antonio, TX, USA, February 23–27, 2013*. 1301–1318.
- [11] Julien Pilourdault, Sihem Amer-Yahia, Senjuti Basu Roy, and Dongwon Lee. 2018. Task Relevance and Diversity as Worker Motivation in Crowdsourcing. In *34th IEEE International Conference on Data Engineering, ICDE 2018, Paris, France, April 16–19, 2018*. 365–376.
- [12] Habibur Rahman, Senjuti Basu Roy, Saravanan Thirumuruganathan, Sihem Amer-Yahia, and Gautam Das. 2019. Optimized group formation for solving collaborative tasks. *VLDB J.* *28*, 1 (2019), 1–23.
- [13] Habibur Rahman, Saravanan Thirumuruganathan, Senjuti Basu Roy, Sihem Amer-Yahia, and Gautam Das. 2015. Worker Skill Estimation in Team-Based Tasks. *Proc. VLDB Endow.* *8*, 11 (2015), 1142–1153.
- [14] Senjuti Basu Roy, Joanna Lykourountzou, Saravanan Thirumuruganathan, Sihem Amer-Yahia, and Gautam Das. 2015. Task assignment optimization in knowledge-intensive crowdsourcing. *VLDB J.* *24*, 4 (2015), 467–491.
- [15] Dong Wei, Senjuti Basu Roy, and Sihem Amer-Yahia. 2020. Recommending Deployment Strategies for Collaborative Tasks. In *Proceedings of the 2020 International Conference on Management of Data, SIGMOD Conference 2020, online conference [Portland, OR, USA], June 14–19, 2020*. 3–17.
- [16] Dong Wei, Senjuti Basu Roy, and Sihem Amer-Yahia. 2020. Task Deployment Recommendation with Worker Availability. In *36th IEEE International Conference on Data Engineering, ICDE 2020, Dallas, TX, USA, April 20–24, 2020*. 1806–1809.