Panel: A Debate on Data and Algorithmic Ethics

Julia Stoyanovich
Drexel University
Philadelphia, PA USA
stoyanovich@drexel.edu

Bill Howe
University of Washington
Seattle, WA USA
billhowe@uw.edu

HV Jagadish
University of Michigan
Ann Arbor, MI USA
jag@umich.edu

Gerome Miklau
University of Massachusetts
Amherst, MA USA
miklau@cs.umass.edu

ABSTRACT

Recently, there has begun a movement towards Fairness, Accountability, and Transparency (FAT) in algorithmic decision making, and in data science more broadly. The database community has not been significantly involved in this movement, despite "owning" the models, languages, and systems that produce the (potentially biased) input to the machine learning applications.

What role should the database community play in this movement? Do the objectives of fairness, accountability and transparency give rise to core data management issues that can drive new research questions and new systems, or are these "soft topics" that are best left to be managed with policy? Will emphasis on these topics dilute our core competency in techniques and technologies for data, or can it reinforce our central role in technology stacks ranging from startups to the enterprise, and from local non-profits to the federal government? The goal of this panel is to debate these questions, and to whet the appetite of the data management community for research in this important emerging area.

PVLDB Reference Format:
Julia Stoyanovich, Bill Howe, HV Jagadish, and Gerome Miklau. Panel: A Debate on Data and Algorithmic Ethics. *PVLDB*, 11 (12): xxxx-yyyy, 2018.
DOI: 10.14778/3229863.3240494

1. INTRODUCTION

As almost everything gets "datafied" and as Big Data has huge impacts on almost every aspect of our lives, it becomes increasingly important to understand the nature of these impacts and to take responsibility for them. In recent history, our technology has produced image labelers [19], search engines [23], and even criminal sentencing systems [4] that discriminate against and denigrate certain races as a byproduct of biased data. In addition to racial bias, examples abound in which algorithmic systems adversely impact members of other historically disadvantaged groups, such as women [10] and persons with disabilities [24]. These are only the high-profile examples that have been exposed through concerted efforts by researchers and computational journalists. These examples are certainly the tip of the iceberg.

Recently, there has begun a movement towards Fairness, Accountability, and Transparency (FAT) in algorithmic decision making [13], and in data science more broadly [21]. As an example of the momentum of this movement, the 2018 FAT* conference had 525 paid registrations and just under 400 people on the waiting list.

The database community has not been significantly involved in this movement, despite "owning" the models, languages, and systems that produce the (potentially biased) input for use with machine learning applications. What role should the database community play in this movement? Do the objectives of fairness, accountability and transparency give rise to core data management issues that can drive new research questions and new systems, or are these "soft topics" that are best left to be managed with policy? Will emphasis on these topics dilute our core competency in data management techniques and technologies, or can it reinforce our central role in technology infrastructure across startups, the enterprise, non-profits, and government?

The goal of this panel is to debate these questions and inform a clear vision for the database community's role. As a side effect of the topic and the debate format, a secondary benefit of the panel is that the audience will become better educated on these issues, which are beginning to dominate the broader data science discussion. We list references to recent work in the next section, and also refer the reader to our recent tutorial [20] and Dagstuhl reports [1, 2].

Proceedings of the VLDB Endowment, Vol. 11, No. 12

DOI: 10.14778/3229863.3240494

2. PANEL STRUCTURE AND DESCRIPTION

The panel is allotted 90 minutes. During the initial 10 minutes, we will explain the structure of the panel to the audience, and will set the stage with a high-level overview of algorithmic ethics, and of the central role that data plays in the algorithmic ethics discourse.

During the next 75 minutes we will debate 3 issues, spending 25 minutes per issue. Each of the three issues will be debated by three panelists: one will serve as the moderator and two as opponents. For each issue, the opponents will present their sides of the argument. Opening arguments will be followed by a moderated round of cross-examination, and, finally, open discussion with questions from the audience. In conclusion, the audience will vote for the winner

among the opponents. The goal of the debate format, with pre-assigned positions, is to explicitly and candidly surface the relevant issues and keep the conversation lively.

During the main part of the panel, we will debate the following three issues:

Is algorithmic and data transparency achievable? Transparency can mean that all code and all data must be made public. However, this transparency interpretation is unrealistic. For example, we may be unable to make all code public due to trade secrets. Perhaps more importantly, making all data public may violate privacy of individuals, and in some cases it may violate laws.

Another aspect to consider is that, even making the code and the data publicly available may not help an end-user understand the process and its effects. This is due to multiple factors, including the conceptual complexity and the subtlety of interactions between the assumptions and design choices and their effects, and the inherent opaqueness of models. Another important factor is the apparent lack of data literacy among the stakeholders: individuals being affected by the processes; human decision makers such as judges, who make decisions with the help of data-driven algorithms; and the general public.

What are the appropriate interpretations of algorithmic and data transparency? What are the technical options for achieving these interpretations? And what are the corresponding database problems that this community can tackle?

This portion of the panel will be informed by recent work on algorithmic and data transparency [3, 8, 9, 15, 18, 26].

What is the "right" definition of algorithmic fairness? We can all agree that algorithmic decision-making should be fair, even if we do not agree on the definition of fairness. But isn't this about algorithm design? Why is this a data problem?

When debating this issue, we will consider the trade-offs between fairness and accuracy, and will discuss recent impossibility results that underscore that different notions of fairness are incompatible. We will also discuss the limits of reasoning about bias and fairness based purely on observational data. Just as in the case of fairness, we can all agree that data should be unbiased, but it can be complicated to determine what exactly it means.

We will pay particular attention to relating these concepts to the stages of the data science lifecycle that are upstream from data analysis: data sharing, cleaning, integration, querying, and ranking.

This portion of the panel will be informed by work on bias in computer systems and on algorithmic fairness, including [5, 6, 7, 11, 12, 14, 16, 17, 22, 25, 27].

Who is responsible? In spite of our best efforts, things will go wrong at times. When mistakes are made, who is to blame? The more complex a system, the more difficult this becomes. For example, suppose the error is due to a mistake in data cleaning. Is the data cleaning algorithm responsible? If the data cleaning algorithm claims that it is operating on a "best effort" basis and does not provide any guarantees of correctness, is the user of the algorithm to blame? Can we afford to use algorithms that do not provide guarantees? To what extent does the answer depend on the context of use (e.g., private vs. public sector), and on what's at stake?

3. PANELISTS

Julia Stoyanovich is an Assistant Professor of Computer Science at Drexel University, and an affiliated faculty at the Center for Information Technology Policy at Princeton. She is a recipient of an NSF CAREER award and of an NSF/CRA CI Fellowship. Julia's research focuses on responsible data management and analysis practices: on operationalizing fairness, diversity, transparency, and data protection in all stages of the data acquisition and processing lifecycle. She established the Data, Responsibly consortium, and serves on the New York City Automated Decision Systems Task Force (by appointment by Mayor de Blasio). In addition to data ethics, Julia works on management and analysis of preference data, and on querying large evolving graphs. She holds M.S. and Ph.D. degrees in Computer Science from Columbia University, and a B.S. in Computer Science and in Mathematics and Statistics from the University of Massachusetts at Amherst.

Bill Howe is Associate Professor in the Information School, Adjunct Associate Professor in Computer Science & Engineering, Senior Data Science Fellow and Founding Associate Director of the UW eScience Institute, Director of the UW Urbanalytics Group, and Founding Chair of the UW Data Science Masters Degree. He has received two Jim Gray Seed Grant awards from Microsoft Research for work on managing scientific data, has had two papers selected for VLDB Journal's Best of Conference issues, and co-authored what are currently the most-cited papers from both VLDB 2010 and ACM SIGMOD 2012. Howe developed a first MOOC on data science that attracted over 200,000 students across two offerings, and founded UW's Data Science for Social Good program. He has a Ph.D. in Computer Science from Portland State University and a Bachelor's degree in Industrial & Systems Engineering from Georgia Tech.

HV Jagadish is the Bernard A Galler Collegiate Professor of Electrical Engineering and Computer Science, and Distinguished Scientist at the Institute for Data Science, at the University of Michigan in Ann Arbor. Prior to 1999, he was Head of the Database Research Department at AT&T Labs, Florham Park, NJ. He is a fellow of the ACM, fellow of AAAS and serves on the board of the Computing Research Association. He has been an Associate Editor for the ACM Transactions on Database Systems (1992-1995), Program Chair of the ACM SIGMOD annual conference (1996), Program Chair of the ISMB conference (2005), a trustee of the VLDB foundation (2004-2009), Founding Editor-in-Chief of the Proceedings of the VLDB Endowment (2008-2014), and Program Chair of the VLDB Conference (2014). Since 2016, he is Editor of the Morgan & Claypool "Synthesis" Lecture Series on Data Management. Among his many awards, he won the ACM SIGMOD Contributions Award in 2013 and the David E Liddle Research Excellence Award (at the University of Michigan) in 2008. He has developed a popular MOOC on Data Science Ethics that is carried by both Coursera and EdX.

Gerome Miklau is a Professor in the College of Information and Computer Sciences at the University of Massachusetts, Amherst. He was an Invited Professor at INRIA

and ENS Cachan for the 2012-2013 academic year. He received the Best Paper Award at the International Conference of Database Theory in 2013, the ACM PODS Alberto O. Mendelzon Test-of-Time Award in 2012, a Lilly Teaching Fellowship in 2011, an NSF CAREER Award in 2007, and he won the 2006 ACM SIGMOD Dissertation Award. He received his Ph.D. in Computer Science from the University of Washington in 2005. He earned Bachelor's degrees in Mathematics and in Rhetoric from the University of California, Berkeley, in 1995.

4. ACKNOWLEDGEMENTS

This work was supported in part by NSF Grants No. 1741047, 1740996, 1741022, and 1741254.

5. REFERENCES

[1] S. Abiteboul, M. Arenas, P. Barceló, M. Bienvenu, D. Calvanese, C. David, R. Hull, E. Hüllermeier, B. Kimelfeld, L. Libkin, W. Martens, T. Milo, F. Murlak, F. Neven, M. Ortiz, T. Schwentick, J. Stoyanovich, J. Su, D. Suciu, V. Vianu, and K. Yi. Research directions for principles of data management (Dagstuhl Perspectives Workshop 16151). *Dagstuhl Manifestos*, 7(1):1–29, 2018.

[2] S. Abiteboul, G. Miklau, J. Stoyanovich, and G. Weikum. Data, Responsibly (Dagstuhl Seminar 16291). *Dagstuhl Reports*, 6(7):42–71, 2016.

[3] P. Adler, C. Falk, S. A. Friedler, T. Nix, G. Rybeck, C. Scheidegger, B. Smith, and S. Venkatasubramanian. Auditing black-box models for indirect influence. *Knowl. Inf. Syst.*, 54(1):95–122, 2018.

[4] J. Angwin, J. Larson, S. Mattu, and L. Kirchner. Machine bias. *ProPublica*, May 23, 2016.

[5] S. Barocas and A. D. Selbst. Big data's disparate impact. *California Law Review*, 104, 2016.

[6] A. Caliskan, J. J. Bryson, and A. Narayanan. Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334):183–186, April 14, 2017.

[7] A. Chouldechova. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *CoRR*, abs/1703.00056, 2017.

[8] D. K. Citron and F. A. Pasquale. The scored society: Due process for automated predictions. *Washington Law Review*, 89, 2014.

[9] A. Datta, S. Sen, and Y. Zick. Algorithmic transparency via quantitative input influence: Theory and experiments with learning systems. In *IEEE Symposium on Security and Privacy, SP 2016*, pages 598–617, 2016.

[10] A. Datta, M. C. Tschantz, and A. Datta. Automated experiments on ad privacy settings. *PoPETs*, 2015(1):92–112, 2015.

[11] C. Dwork, M. Hardt, T. Pitassi, O. Reingold, and R. S. Zemel. Fairness through awareness. In *Innovations in Theoretical Computer Science 2012, Cambridge, MA, USA, January 8-10, 2012*, pages 214–226, 2012.

[12] S. A. Friedler, C. Scheidegger, and S. Venkatasubramanian. On the (im)possibility of fairness. *CoRR*, abs/1609.07236, 2016.

[13] S. A. Friedler and C. Wilson, editors. *Conference on Fairness, Accountability and Transparency, FAT 2018*, volume 81 of *Proceedings of Machine Learning Research*. PMLR, 2018.

[14] B. Friedman and H. Nissenbaum. Bias in computer systems. *ACM Trans. Inf. Syst.*, 14(3):330–347, 1996.

[15] S. Galhotra, Y. Brun, and A. Meliou. Fairness testing: testing software for discrimination. In *Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering, ESEC/FSE 2017, Paderborn, Germany, September 4-8, 2017*, pages 498–510, 2017.

[16] M. Hardt, E. Price, and N. Srebro. Equality of opportunity in supervised learning. *CoRR*, abs/1610.02413, 2016.

[17] J. M. Kleinberg, S. Mullainathan, and M. Raghavan. Inherent trade-offs in the fair determination of risk scores. In *8th Innovations in Theoretical Computer Science Conference, ITCS 2017, January 9-11, 2017, Berkeley, CA, USA*, pages 43:1–43:23, 2017.

[18] J. A. Kroll, J. Huey, S. Barocas, E. W. Felten, J. R. Reidenberg, D. G. Robinson, and H. Yu. Accountable algorithms. *University of Pennsylvania Law Review*, 165, 2017.

[19] R. Nieva. Google apologizes for algorithm mistakenly calling black people 'gorillas'. *CNet*, July 4, 2015.

[20] J. Stoyanovich, S. Abiteboul, and G. Miklau. Data responsibly: Fairness, neutrality and transparency in data analysis. In *Proceedings of the 19th International Conference on Extending Database Technology, EDBT 2016, Bordeaux, France, March 15-16, 2016, Bordeaux, France, March 15-16, 2016.*, pages 718–719, 2016.

[21] J. Stoyanovich, B. Howe, S. Abiteboul, G. Miklau, A. Sahuguet, and G. Weikum. Fides: Towards a platform for responsible data science. In *Proceedings of the 29th International Conference on Scientific and Statistical Database Management, Chicago, IL, USA, June 27-29, 2017*, pages 26:1–26:6, 2017.

[22] J. Stoyanovich, K. Yang, and H. V. Jagadish. Online set selection with fairness and diversity constraints. In *Proceedings of the 21th International Conference on Extending Database Technology, EDBT 2018, Vienna, Austria, March 26-29, 2018.*, pages 241–252, 2018.

[23] L. Sweeney. Discrimination in online ad delivery. *Commun. ACM*, 56(5):44–54, 2013.

[24] L. Weber and E. Dwoskin. Are workplace personality tests fair? *Wall Street Journal*, September 29, 2014.

[25] K. Yang and J. Stoyanovich. Measuring fairness in ranked outputs. In *Proceedings of the 29th International Conference on Scientific and Statistical Database Management, Chicago, IL, USA, June 27-29, 2017*, pages 22:1–22:6, 2017.

[26] K. Yang, J. Stoyanovich, A. Asudeh, B. Howe, H. Jagadish, and G. Miklau. A nutritional label for rankings. In *Proceedings of the 2018 ACM SIGMOD International Conference on Management of Data*, 2018.

[27] I. Zliobaite. Measuring discrimination in algorithmic decision making. *Data Min. Knowl. Discov.*, 31(4):1060–1089, 2017.