Patterns

CellPress
OPEN ACCESS

People of Data

Two researchers share how their cross disciplinary collaboration enables work to guide the future of data science

Maggie Delano[1,*] and Kendra Albert[2]
[1]Swarthmore College, Swarthmore, PA, USA
[2]Harvard Law School, Cambridge, MA, USA
*Correspondence: mdelano1@swarthmore.edu
https://doi.org/10.1016/j.patter.2022.100573

In their recent perspective published in *Patterns*, Maggie Delano and Kendra Albert highlight the limitations of sex and gender data classification in health systems and show how this contributes to the marginalization of trans and non-binary individuals. They provide recommendations to improve incorporating gender data into healthcare algorithms. Here they discuss their collaboration and how it enabled this cross-disciplinary research.

What would you like to share about your background (personal and/or professional)?

Maggie Delano: My background is in electrical engineering, particularly analog circuit design and embedded systems. My research has focused on using these skills to develop wearable and portable monitoring systems for biomedical applications, specifically cardiovascular monitoring. My master's thesis focused on building a wearable electrocardiogram (ECG) monitoring system, and my PhD thesis focused on building and testing a portable fluid status monitor for fluid overload management. I also have a background in gender studies and inclusive engineering design, which were particularly relevant for our recent paper published in *Patterns*.[1] During the pandemic, I started working with a group of researchers on human subjects testing for adversarial machine learning applications, leveraging what I had learned about testing biomedical devices. We learned a lot about where the machine learning field was with respect to testing and validation, which was part of the inspiration for the current paper we wrote.

Kendra Albert: I have a relatively unusual background for this kind of academic work—my day job is as a technology lawyer, where I teach students to practice law by working with real clients, but I also teach and do research in transgender studies. Both of these areas are inherently interdisciplinary and require being able to understand the impacts of broad structural interventions on people's lives and choices, which is something that shows up across all of my academic work. The work with Maggie comes very much out of my own personal experience as a non-binary person who regularly encounters discriminatory or exclusionary technology. When we started doing the work, I expected there would already be a robust field of writing in this area, and when there wasn't, I felt that we could really help others contextualize their experiences by providing frameworks and terms.

What motivated you to become a (data) researcher? Is there anyone/anything that helped guide you on your path?

MD: My background is in building devices to monitor physiological signals about the body, which is really all about taking good measurements. Unfortunately, a major theme of our current paper is that many data scientists take this first step for granted. I realized that I can make important contributions as a data researcher by helping data scientists understand where their data are coming from and what a dataset's limitations are. I learned to be cautious in interpreting data from my PhD advisor Charlie Sodini. Whenever we presented a figure during our group meetings, Charlie would make us clearly articulate the values and units on each axis and explain why the results looked the way that they did. If we didn't know, we needed to have a plan to figure it out. There was no hand waving or ignoring outliers. This helped me be much more rigorous in interpreting data, taking my time to really understand what is going on.

KA: It's funny; I don't think I would consider myself a data researcher! But in terms of research, I know how helpful examples and theoretical frames have been for processing my own experiences and understanding the world, so I try to produce research that creates that framework for other people. Also, reading is fun.

What is the definition of data science in your opinion? What is a data scientist? Do you self-identify as one?

MD: I think the field of data science has emerged out of a recognition that researchers across disciplines are collecting so much data that there's space to specialize in its manipulation and interpretation. Since my background and specialty is in engineering design and measurement, I don't personally identify as a data scientist, but I see my work as operating at an earlier stage, generating data that data scientists can draw insights from. I can leverage what I know about measurement to help data scientists better understand how to use it.

KA: Yeah, I don't self-identify as a data scientist! I'll leave the data science conversation to the data scientists.

Why did you decide to publish in *Patterns*?

We were approached by an editor at *Patterns* after we wrote an article about sex/gender and weight scales for the

Kendra Albert (left) and Maggie Delano (right)

Fairness, Accountability and Transparency Conference in 2021.[2] The editor invited us to expand upon that paper, and we decided we wanted to spend time focusing on how sex/gender are used in machine learning models broadly.

How did this project you wrote about come to be?

After being approached by an editor to publish a follow up paper in *Patterns*, we decided we were interested in how sex/gender is embedded in the medical system more broadly beyond just weight scales and body-composition analysis. We had worked with a few other researchers to submit papers to machine learning conferences about human subjects testing and broader impact statements, and we thought it would be interesting to focus specifically on the intersection of medicine, machine learning, and sex/gender.

Although the paper does talk about machine learning contexts specifically, we hope that the phenomena we name are useful to other fields as well!

How important was the collaboration to the success of the paper?

Collaboration was essential for the success of this paper. In addition to dividing up the work, collaborating allowed us to bounce ideas off each other and build on our relative strengths. Maggie brought their background knowledge in pathophysiology to the work, and Kendra brought their experience in transgender studies. Maggie focused on understanding background literature and the history of electronic health records (EHRs) and sex/gender data. They also focused on the recommendations section to provide suggestions for researchers. Kendra was then able to focus on a deep dive into medical machine learning, including conducting the research for the case study in the article.

How important do you think collaboration is in general to research?

We think collaboration is almost required in an area like the one we are working in. We see our work at the intersection of data science/machine learning, medicine, and gender studies. Few people are an expert in all of these fields and their subfields that might be relevant to the work; by working together we were able to cover more ground than if only one of us were writing this article. Having co-authors with different backgrounds is also really helpful to get a sense of what is "obvious" and what is worth writing about. Things might seem clear to experts in a subfield that researchers in another subfield could benefit from learning about, but without collaboration these cross-disciplinary connections can be hard to identify. Our earlier work, "This Whole Thing Smacks of Gender,"[2] is very much an example of how things that are obvious to a person in one space may be news to a person in another and how those areas can represent fruitful places to focus.

REFERENCES

1. Albert, K., and Delano, M. (2022). Sex trouble: sex/gender slippage, sex confusion, and sex obsession in machine learning using electronic health records. Patterns *3*. https://doi.org/10.1016/j.patter.2022.100534.
2. Albert, K., and Delano, M. (2021). This Whole Thing Smacks of Gender: Algorithmic Exclusion in Bioimpedance-based Body Composition Analysis. In Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (Association for Computing Machinery), pp. 342–352. https://doi.org/10.1145/3442188.3445898.

About the authors

Maggie Delano is an assistant professor at Swarthmore College, where they teach computer engineering and inclusive engineering design. Their research focuses on the development of a wearable monitoring system for patients with congestive heart failure. They also write about and research better design, including the use of sex/gender in medicine and human subjects testing in machine learning. More details about Maggie can be found on their website: https://www.maggiedelano.com/

Kendra Albert is a public interest technology lawyer whose research interests include gender, machine learning, and power. They serve as a clinical instructor at the Cyberlaw Clinic at Harvard Law School, where they teach students to practice law by working with *pro bono* clients. Kendra also has taught classes on transgender rights in the Program on Studies of Women, Gender, and Sexuality at Harvard College and Harvard Law School. More at https://kendraalbert.com/.