




# Unfairness Is Everywhere, so What to Do? An Interview With Jeanna Matthews

Brittany Johnson and Tim Menzies 

## From the Editors

Usually, when we talk to other software engineers about fairness and discrimination, it quickly becomes a conversation about measurement (e.g., how to check if different populations within society are getting different false-positive rates from that software). But if you talk to Dr. Jeanna Mathews from Clarkson University, the conversation is very different. She focuses on risk as a function of the severity of consequences and the probability of those consequences occurring. Then she asks how legislation could help manage high-risk software.

And for future issues we ask, “What do you want to see in this ‘SE for Ethics’ column”? Do you have an important insight or industrial case study? Something that could prompt an important discussion? Or, alternatively, that extends or challenges significant ideas? If so, email a one-paragraph synopsis to [johnsonb@gmu.edu](mailto:johnsonb@gmu.edu) or [tim@ieee.org](mailto:tim@ieee.org) (subject line: “SE for Ethics: Idea: [Your Idea]”). If that looks interesting, we’ll ask you to submit a 1,000- to 3,000-word article (where each graph, table, or figure is worth 250 words) for review for *IEEE Software*.—Tim Menzies

**“UNFAIRNESS IS EVERYWHERE,”** says Dr. Jeanna Matthews from Clarkson University. “We say we want a fair society, but we are far

from a level playing field. If we ask ‘what is the secret of your success?’, people often think of their own hard work or a good decision they made. However, It is often more accurate to look at advantages, like the ability to borrow money from family and

friends when you are in trouble, deep network connections so that you hear about opportunities or have a human look at your application, the ability to move on from a mistake that might send someone else to jail, help at home to care for children, etc.

Digital Object Identifier 10.1109/MS.2023.3305722  
Date of current version: 1 December 2023

## ABOUT DR. JEANNA MATTHEWS



Dr. Jeanna Matthews, professor of computer science at Clarkson University, Potsdam, NY 13699 USA, earned her Ph.D. C.S. from the University of California, Berkeley, in 1999. She is a founding Chair of the Association for Computing Machinery (ACM) Technology Policy Subcommittee on Artificial Intelligence and Algorithmic Accountability, an ACM Distinguished Speaker, and a Fulbright Scholar. Her current work focuses on securing societal decision-making processes and supporting the rights of individuals in a world of automation. For more information, see <https://people.clarkson.edu/~jmatthew/>.



Dr. Jeanna Matthews.

software are often deemed admissible in court using articles in peer-reviewed publications as evidence of acceptability. But peer review of an article is a long way from thorough verification and validation of the underlying system. Peer reviewers are answering whether the results are interesting to the research community, not whether the software is sufficiently reliable. There has been reluctance to hold criminal justice software to the same rigorous standards for verification and validation as software used in other areas, such as medical devices or air traffic control for example.” Of special interest to software engineers, there has been a surprising reluctance to include software engineering in the relevant scientific community of interest when determining whether the output of a particular piece of software had sufficiently gained general acceptance for admissibility as evidence according to the Frye and Daubert standards.”<sup>8</sup>

The narrative that success comes from hard work misses that many people work hard and never succeed. Success often comes from exploiting a playing field that is far from level and, when push comes to shove, we often want those advantages for our children, our family, our friends, our community, our organizations.”

It is hardly surprising that this lack of a level playing field is reflected in our software too. For example, Amazon had to scrap an automated recruiting tool as it was found to be biased against women.<sup>1</sup> A widely used face recognition software was found to be biased against dark-skinned women.<sup>2</sup> Google Translate, the most popular translation engine in the world, was shown to exhibit gender bias. “She is an engineer, He is a nurse” translated into Turkish and then again into English became “He is an engineer, She is a nurse.”<sup>3</sup> The Compas recidivism prediction model, used in courts across America, is more likely to mistakenly

predict that black defendants are high risk, while making the opposite type of mistake for white defendants.<sup>4</sup> And every day this list grows longer and longer.

### Criminal Justice Software

Since unfairness from software is potentially everywhere, Dr. Matthews works to manage and mitigate its effects. Some of her research concerns the software used in the criminal justice systems.<sup>5,6,7</sup> “In a democracy, we believe in a fair trial. We say ‘innocent until proven guilty’ and that we have the ‘right to confront our accusers.’ But what does that mean? If software generates evidence that could lead to someone getting convicted, and they can’t question that software, and that software is not held to a high standard of accuracy, are those principles we think are following still even true?”

She reports many issues with the software used in the criminal justice system. “Results of forensics

### Industrial Regulation

Policy makers around the world are debating regulatory approaches to control automated systems, especially in response to growing concern over generative artificial intelligence technologies, like ChatGPT and DALL-E. Matthews has been involved in crafting recommendations for technology policy for many years. She recommends a laser focus on risk. “Some regulatory proposals focus on specific technologies, like large language models or facial recognition. Some make distinctions between large and small companies or between application areas. I recommend a laser focus on risk as defined by the severity of potential consequences and the probability of those consequences.” She explains that

## ABOUT THE AUTHORS



**BRITTANY JOHNSON** is an assistant professor in computer science at George Mason University, Fairfax, VA 22030 USA. Contact her at [johnsonb@gmu.edu](mailto:johnsonb@gmu.edu).



**TIM MENZIES** is a full professor at North Carolina State University, Raleigh, NC 27606 USA. Contact him at [timmm@ieee.org](mailto:timmm@ieee.org).

for everyone. That we are making the same decisions and taking the same actions we always have, just faster and more efficiently. In my experience, this is rarely the case. We are changing the fundamental nature of the decisions being made and actions being taken. There will be new winners and losers. Software is putting its thumb on the scale. I want transparency so people can understand the tradeoffs being made and advocate better for themselves: as citizens, as customers, as workers, as individuals, as members of many different groups. When we don't recognize the tradeoffs being made, someone else will be making those decisions in their own best interest, not ours."

focus on risk can help manage the most severe risks without dampening innovation.

Mathews recommends that policy makers take advice from long-standing software engineering principles, like those found in the IEEE 1012 standard.<sup>9</sup> "The IEEE 1012 standard does a nice job of assigning different integrity levels or risk tiers to systems based on the severity and probability of possible consequences and assigning more verification and validation requirements to higher risk tiers. For example, voluntary self-policing standards might be fine for low-risk software, but for high-risk software, we could require substantial risk management and independent verification and validation."

She also recommends involving stakeholders broadly defined throughout the lifecycle of software. "Organizations will have incentives to manage risks to themselves and their customers, but risk management really needs to consider those

impacted by the system, society as a whole, the environment, etc. I think we are going to need more than voluntary compliance with best practices to accomplish this. For high-risk software, stakeholders broadly defined should be involved from concept to design, implementation, deployment, and beyond. Even if an organization estimates the risks to be low, they should still be tracking and reporting actual consequences after deployment and adjusting based on evidence. It wouldn't be inappropriate to have penalties for estimating the system to be low risk in the first place when that turns out not to be true. Many risks are foreseeable and manageable if you involve stakeholders broadly defined in the process of risk management from the beginning."

### Transparency and Access

For Dr. Mathews, it's all about transparency. "We often tell the story that automation will be great

### References

1. "Amazon scraps secret AI recruiting tool that showed bias against women," *Reuters*. [Online]. Available: <https://www.reuters.com/article/us-amazoncom-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-thatshowed-bias-against-women-idUSKCN1MK08G>
2. "Study finds gender and skin-type bias in commercial artificial-intelligence systems," *MIT News*, 2018. [Online]. Available: <http://news.mit.edu/2018/study-finds-genderskin-type-bias-artificial-intelligence-systems-0212>
3. A. Caliskan, J. J. Bryson, and A. Narayanan, "Semantics derived automatically from language corpora contain human-like biases," *Science*, vol. 356, no. 6334, pp. 183–186, Apr. 2017, doi: 10.1126/science.aal4230. [Online]. Available: <https://science.sciencemag.org/content/356/6334/183>
4. C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nature*

*Mach. Intell.*, vol. 1, no. 5, pp. 206–215, May 2019, doi: 10.1038/s42256-019-0048-x.

5. J. Matthews et al., “The right to confront your accusers: Opening the black box of forensic DNA software,” in *Proc. AAAI/ACM Conf. AI, Ethics, Soc. (AIES)*, New York, NY, USA: Association for Computing Machinery, 2019, pp. 321–327, doi: 10.1145/3306618.3314279.
6. J. N. Matthews et al., “When trusted black boxes don’t agree: Incentivizing iterative improvement and accountability in critical software systems,” in *Proc. AAAI/ACM Conf. AI, Ethics, Soc. (AIES)*, New York, NY, USA: Association for Computing Machinery, 2020, pp. 102–108, doi: 10.1145/3375627.3375807.
7. J. Matthews, B. Hedin, and M. Canellas, “Trustworthy evidence for trustworthy technology: An overview of evidence for assessing the trustworthiness of autonomous and intelligent systems,” IEEE-USA, Washington, DC, USA, Sep. 2022. [Online]. Available: [https://ieeusa.org/assets/public-policy/committees/aipc/IEEE\\_Trustworthy-Evidence-for-Trustworthy-Technology\\_Sept22.pdf](https://ieeusa.org/assets/public-policy/committees/aipc/IEEE_Trustworthy-Evidence-for-Trustworthy-Technology_Sept22.pdf)
8. M. Canellas, “Defending IEEE software standards in federal criminal court,” *Computer*, vol. 54, no. 6, pp. 14–23, Jun. 2021, doi: 10.1109/MC.2020.3038630.
9. J. Matthews, “How should we regulate AI? Practical strategies for regulation and risk management from the IEEE 1012 standard for system, software, and hardware verification and validation,” IEEE-USA, Washington, DC, USA, Aug. 2023. [Online]. Available: <https://ieeusa.org/assets/public-policy/committees/aipc/How-Should-We-Regulate-AI.pdf>

**SUBMIT  
TODAY**

IEEE TRANSACTIONS ON

**SUSTAINABLE COMPUTING**

**SUBSCRIBE AND SUBMIT**

For more information on paper submission, featured articles, calls for papers, and subscription links visit: [www.computer.org/tsusc](http://www.computer.org/tsusc)



Digital Object Identifier 10.1109/MS.2023.3328468