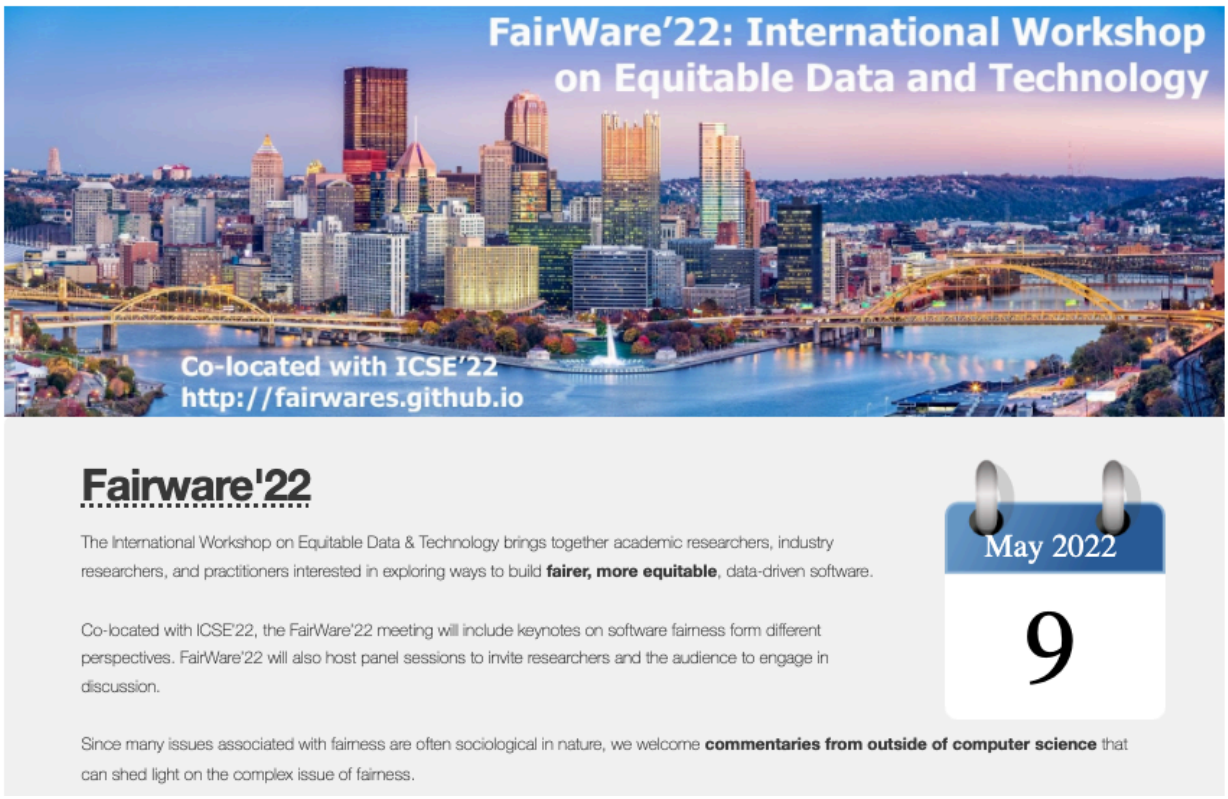

RESOURCES

SOFTWARE

- https://aif360.mybluemix.net/
- (Early from bachelor thesis) Tool to visualize (different) bias in word embeddings: https://github.com/NiklasFinzel/EmbeddingBiasTool

PAPERS:

- Motivation to get us working harder:
  - Kate Starbird telling anyone saying “we should not try” https://www.youtube.com/embed/9gzo-1jK-TA
  - Geraldine Fitzpatrixk’s high impact on design decisions participant Thomas:
  - http://superflux.in/index.php/work/uninvited-guests/#
- Microsoft Responsible AI principles in practice
  - https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1:primaryr6
- Diversity May Be Key to Reducing Errors in Quantum Computing: In quantum computing, as in team building, a little diversity can help get the job done better, computer scientists have found https://news.gatech.edu/news/2019/10/14/diversity-may-be-key-reducing-errors-quantum-computing
- The more we learn about each other, the less we are biased. Richer features rule! Bias arises when a model“reaches too far” towards some outlier region that was rarely encountered in training Therefore: Bias can be removed via extrapolation that smooths away those outlier points. https://arxiv.org/pdf/2110.01109.pdf
- MozFest Spelman College Blackpaper
  - Black Women’s Experiences with Algorithmic Microaggressions
  - https://foundation.mozilla.org/en/insights/mozfest-spelman-college-blackpaper/
- Deconstructing Community-Based Collaborative Design: Towards More Equitable Participatory Design Engagements
  - https://dl.acm.org/doi/pdf/10.1145/3359318
  - CHRISTINA N. HARRINGTON, Northwestern University, USA
  - SHEENA ERETE, DePaul University, USA
  - ANNE MARIE PIPER, Northwestern University, USA

- NIST Internal Report 8351-DRAFT DNA Mixture Interpretation: A NIST Scientific Foundation Review
- publications from SE on fairness testing
  - 2017 Fairness Testing: Testing Software for Discrimination Sainyam Galhotra Yuriy Brun Alexandra Meliou https://people.cs.umass.edu/~brun/pubs/pubs/Galhotra17fse.pdf
  - Bias in Machine Learning Software: Why? How? What to Do? Joymallya Chakraborty, Suvodeep Majumder, Tim Menzies https://arxiv.org/pdf/2105.12195.pdf

- Publications that use fairness in the splitting criteria of Decision Trees:
  - -Discrimination aware decision tree learning
    - https://www.researchgate.net/profile/Mykola-Pechenizkiy/publication/220765322_Discrimination_Aware_Decision_Tree_Learning/links/0deec5304e7cd38164000000/Discrimination-Aware-Decision-Tree-Learning.pdf
  - -Faht: an adaptive fairness-aware decision tree classifier
    - https://www.ijcai.org/proceedings/2019/0205.pdf

- https://lin-web.clarkson.edu/~jmatthew/publications/NJvPickett_202010.pdf
- The Right To Confront Your Accusers: Opening the Black Box of Forensic DNA Software

- https://dl.acm.org/citation.cfm?id=3314279
- When Trusted Black Boxes Don't Agree: Incentivizing Iterative Improvement and Accountability in Critical Software Systems
- https://dl.acm.org/doi/abs/10.1145/3375627.3375807

- There are 1000s of people incarcerated on evident that we now view to be unreliable
    - https://www.pbwt.com/second-circuit-blog/second-circuit-oks-use-of-now-defunct-dna-testing-method
    - https://themarkup.org/news/2021/03/09/powerful-dna-software-used-in-hundreds-of-criminal-cases-faces-new-scrutiny
    - https://www.nytimes.com/2017/09/04/nyregion/dna-analysis-evidence-new-york-disputed-techniques.html
    - Amicus Brief https://lin-web.clarkson.edu/~jmatthew/publications/NJvPickett_202010.pdf

- https://ieeeusa.org/committees/aipc/
- Resources
    - https://lin-web.clarkson.edu/~jmatthew/publications/Fairware_Panel_20220509.html

- FYI - as mentioned earlier here is a link to my draft (pre-publication) report on "The Future of Conferences" post-pandemic - based on a recent ACM, IEEE, Agile Alliance community survey. Feedback welcome. Cheers, Stev
    - https://manclswx.com/papers/Future%20of%20Conference%20Research%20Survey%20Report%20May%202022.pdf

READ THIS:

1. Jie Zhang's latest paper "survey paper bridge AI and fairness"
2. Value sensitive design. Shaping technology with normal imagination
3. https://katta.mere.st/research/publications/
4. https://lin-web.clarkson.edu/~jmatthew/films2021/
5. Ben Green. Flaws of policies of requiring human oversight
6. The Intersectional Environmentalist– Leah Thomas
7. Techno-Vernacular Creativity and Innovation - Nettrice Gasking The Intersectional https://arxiv.org/pdf/2109.05067
8. Algorithms of oppression
    a. https://nyupress.org/9781479837243/algorithms-of-oppression/
9. The intersectionality wars When Kimberlé Crenshaw coined the term 30 years ago, it was a relatively obscure legal concept. Then it went viral. By Jane Coaston
    a. https://www.vox.com/the-highlight/2019/5/20/18542843/intersectionality-conservatism-law-race-gender-discrimination

ABSTRACTS

The Flaws of Policies Requiring Human Oversight of Government Algorithms

Ben Green

As algorithms become an influential component of government decision-making around the world, policymakers have debated how governments can attain the benefits of algorithms while preventing the harms of algorithms. One mechanism that has become a centerpiece of global efforts to regulate government algorithms is to require human oversight of algorithmic decisions. Despite the widespread turn to human oversight, these policies rest on an uninterrogated assumption: that people are able to effectively oversee algorithmic decision-making. In this article, I survey 41 policies that prescribe human oversight of government algorithms and find that they suffer from two significant flaws. First, evidence suggests that people are unable to perform the desired oversight functions. Second, as a result of the first flaw, human oversight policies legitimize government uses of faulty and controversial algorithms without addressing the fundamental issues with these tools. Thus, rather than protect against the potential harms of algorithmic decision-making in government, human oversight policies provide a false sense of security in adopting algorithms and enable vendors and agencies to shirk accountability for algorithmic harms. In light of these flaws, I propose a shift from human oversight to institutional oversight as the central mechanism for regulating government algorithms. This institutional approach operates in two stages. First, agencies must justify that it is appropriate to incorporate an algorithm into decision-making and that any proposed forms of human oversight are supported by empirical evidence. Second, these justifications must receive democratic public review and approval before the agency can adopt the algorithm.

https://arxiv.org/pdf/2109.05067

Humans in the Loop

Crootof, Rebecca and Kaminski, Margot E. and Price II, William Nicholson, **Humans in the Loop** (March 25, 2022). Vanderbilt Law Review, Forthcoming 2023, U of Colorado Law Legal Studies Research Paper No. 22-10, U of Michigan Public Law Research Paper No. 22-011, Available at SSRN: https://ssrn.com/abstract=4066781 or http://dx.doi.org/10.2139/ssrn.4066781

Abstract:

From lethal drones to cancer diagnostics, complex and artificially intelligent algorithms are increasingly integrated into decisionmaking that affects human lives, raising challenging questions about the proper allocation of decisional authority between humans and machines. Regulators commonly respond to these concerns by putting a “human in the loop”: using law to require or encourage including an individual within an algorithmic decisionmaking process.

Drawing on our distinctive areas of expertise with algorithmic systems, we take a bird’s eye view to make three generalizable contributions to the discourse. First, contrary to the popular narrative, the law is already profoundly (and problematically) involved in governing algorithmic systems. Law may explicitly require or prohibit human involvement and law may indirectly encourage or discourage human involvement, all without regard to what we know about the strengths and weaknesses of human and algorithmic decisionmakers and the particular quirks of hybrid human-machine systems. Second, we identify “the MABA-MABA trap,” wherein regulators are tempted to address a panoply of concerns by

“slapping a human in it” based on presumptions about what humans and algorithms are respectively better at doing, often without realizing that the new hybrid system needs its own distinct regulatory interventions. Instead, we suggest that regulators should focus on what they want the human to do—what role the human is meant to play—and design regulations to allow humans to play these roles successfully. Third, borrowing concepts from systems engineering and existing law regulating railroads, nuclear reactors, and medical devices, we highlight lessons for regulating humans in the loop as well as alternative means of regulating human-machine systems going forward.

NOTES FROM CHAT STREAM

CHI 2021, more workshops and papers that design with social justice at center (Angela may have specific resources to add)

One example from 2011 re ability-based design speaking to the accessibility community: https://dl.acm.org/doi/10.1145/1952383.1952384

Participatory design in research with vulnerable populations (Angela's course)

Jaye Nias "black paper"
https://foundation.mozilla.org/en/insights/mozfest-spelman-college-blackpaper/

"Bias arises when a model"reaches too far" towards some
outlier region that was rarely encountered in training.
Therefore:
Bias can be removed via extrapolation that smooths
away those outlier points."
https://arxiv.org/pdf/2110.01109.pdf

https://www.news.gatech.edu/2019/10/14/diversity-may-be-key-reducing-errors-quantum-computing

link to my draft (pre-publication) report on "The Future of Conferences" post-pandemic - based on a recent ACM, IEEE, Agile Alliance community survey. Feedback welcome. Cheers, Steve

https://manclswx.com/papers/Future%20of%20Conference%20Research%20Survey%20Report%20May%202022.pdf

A awareness
B behavior
C culture
- Geraldine Fitzpatrick
http://superflux.in/index.php/work/uninvited-guests/#

"There's power in stories"

Community-based data collection: https://wemeasure.org/tools-services/

Must-reads (beyond algorithmic fairness):
https://katta.mere.st/research/publications/
The Intersectional Environmentalist

IDEAS FOR 2023

- More geographic diversity!!!
- Bring more developers
- Jaye: non-reviewed as inspiring
- +1 on tutorial, that could be a good option to spread some of the awesome ongoing work that has been done in this area.
- Solicit different contributions. Beyond paper formats.
    - Videos , extended abstracts? Ideas for panel.?
    - Experience reports with shepherding?
    - Rebecca worths-brook?
- i loved the legal participation. please invite them again!!!!
- Communities in the education space. Features assignments?
- Open space, unconference format. ??
- Grass root organizations can inspire us.??
- Debate contrarian opinion : show trial of gang of 4??