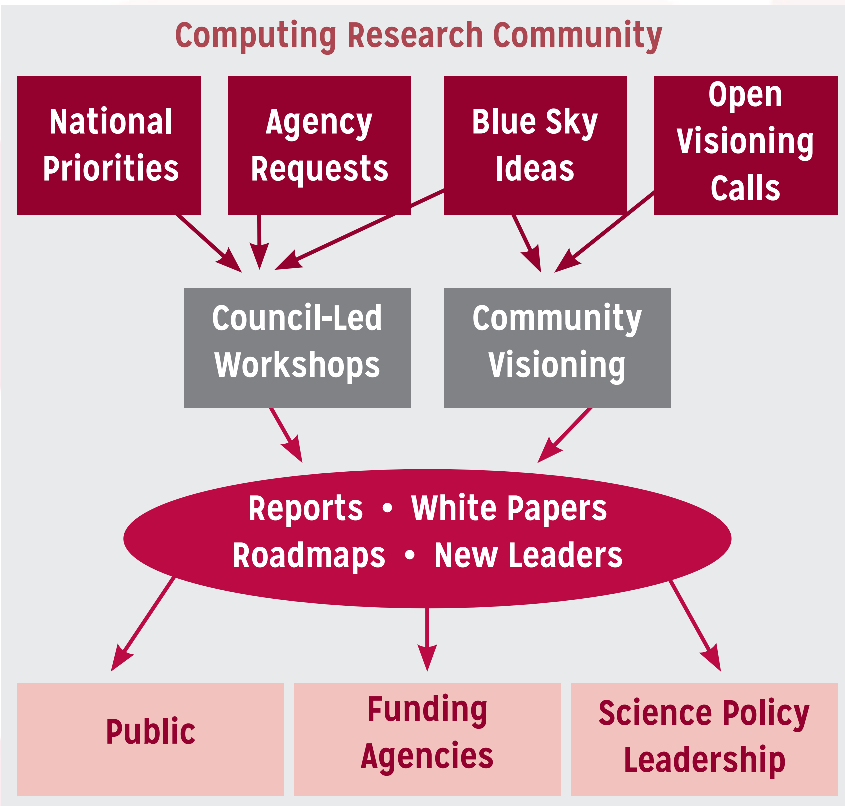
COMPUTING COMMUNITY CONSORTIUM

The **mission** of the Computing Research Association's Computing Community Consortium (CCC) is to **catalyze** the computing research community and **enable** the pursuit of innovative, high-impact research.



Bring the computing research community together to envision audacious research challenges.

Communicate these challenges and opportunities to the broader national community.

Facilitate investment in these research challenges **by key stakeholders**.

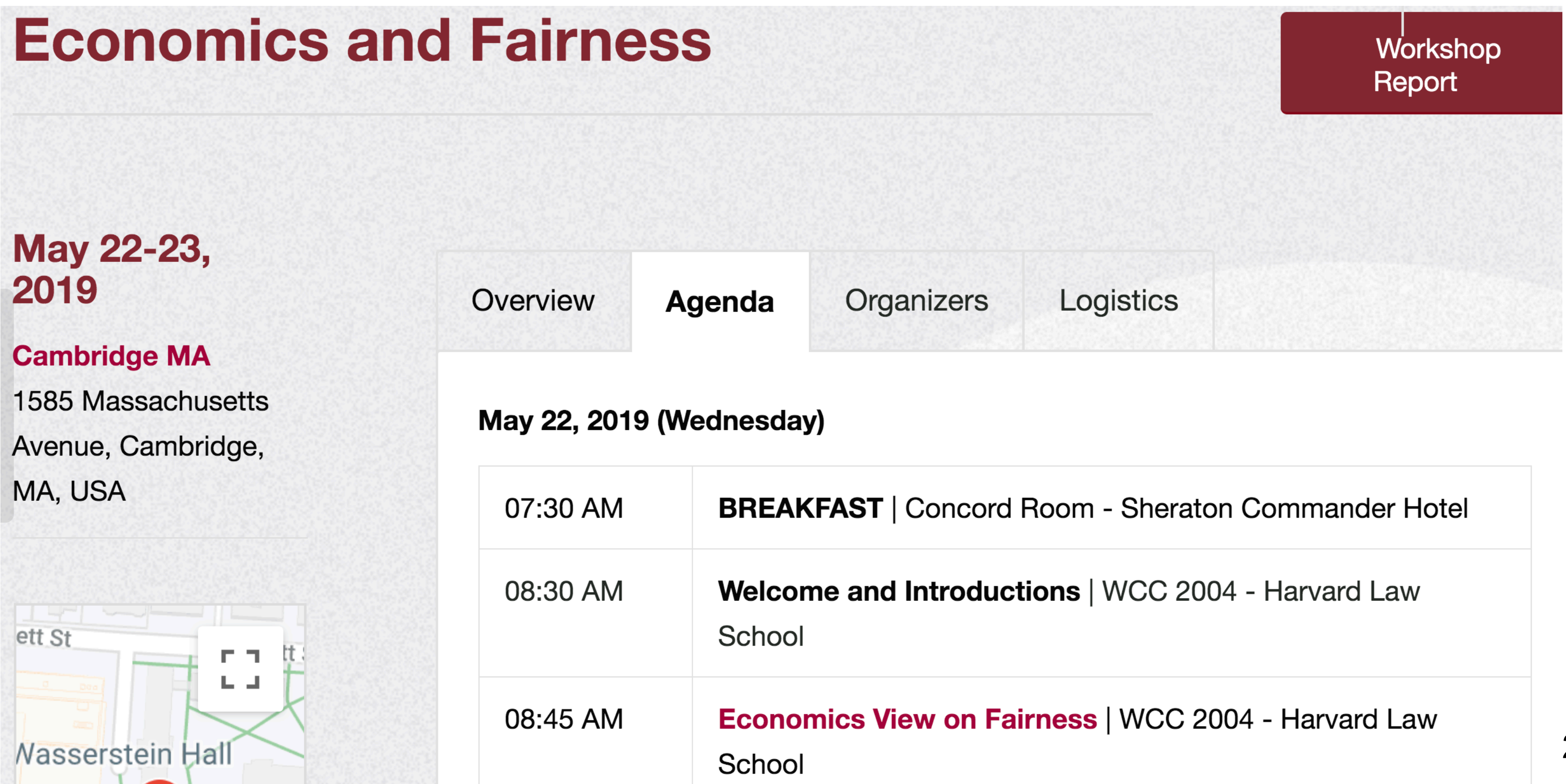
Inculcate values of **leadership** and service by the computing research community.

Inform and influence early career researchers to engage in these community-led research challenges.

Visioning workshop: Algorithmic and Economic Perspectives on Fairness

Co-chairs: David Parkes (Harvard), Rakesh Vohra (Penn)

CCC Fairness and Accountability Task force: Liz Bradley, Sampath Kannan, Ronitt Rubinfeld, David Parkes, Suresh Venkatasubramanian


Economics and Fairness

Workshop Report

May 22-23, 2019

Cambridge MA

1585 Massachusetts Avenue, Cambridge, MA, USA

Overview | **Agenda** | Organizers | Logistics

May 22, 2019 (Wednesday)

07:30 AM	**BREAKFAST** | Concord Room - Sheraton Commander Hotel
08:30 AM	**Welcome and Introductions** | WCC 2004 - Harvard Law School
08:45 AM	**Economics View on Fairness** | WCC 2004 - Harvard Law School

- The workshop discussed **methods to ensure economic fairness in a data-driven world**. Participants were asked to identify and frame what they thought were the most pressing issues and outline concrete problems

08:30 AM	**Welcome and Introductions** | WCC 2004 - Harvard Law School
08:45 AM	**Economics View on Fairness** | WCC 2004 - Harvard Law School Mallesh Pai, Rice- "Can Free Markets lead to Fair Markets?" Recommended Reading: • The Economics of Discrimination by Gary Becker • Theories of Statistical Discrimination and Affirmative Action: A Survey by Hanming Fang and Andrea Moro

09:30 AM	**Computer Science View on Fairness** | WCC 2004 - Harvard Law School Sharad Goel, Stanford– "The Measure and Mismeasure of Fairness: A Critical Review of Fair Machine Learning" Recommended Reading: • The Measure and Mismeasure of Fairness: A Critical Review of Fair Machine Learning* by Sam Corbett-Davis and Sharad Goel

10:30 AM	**Algorithm Decision Making** | WCC 2004 - Harvard Law School Prasanna Tambe, Wharton – "Artificial Intelligence in Human Resources Management: Challenges and a Path Forward" Recommended Reading: • Artificial Intelligence in Human Resources Management: Challenges and a Path Forward by Prasanna Tambe Lindsey Zuloaga, HireVue – "Algorithms for Hire" Bo Cowgill, Columbia – "Economics, Fairness and Algorithmic Bias" Recommended Reading: • Bias and Productivity in Humans and Machines by Bo Cowgill • Economics, Fairness and Algorithmic Bias by Bo Cowgill and Catherine E. Tucker Discussant: Matt Weinberg, Princeton

03:00 PM	**Platforms** | WCC 2004 - Harvard Law School Daniel Knoefle, Uber- “Pricing Efficiently in Designed Markets: Evidence from Ride-Sharing” Recommended Reading: • Pricing Efficiently in Designed Markets: Evidence from Ride-Sharing by Jonathan Hall, John Horton, Daniel Knoefle Karen Levy, Cornell University- “Trade-offs in Designing Against Discrimination” Recommended Reading: • Designing Against Discrimination in Online Markets by Karen Levy and Solon Barocas Mike Luca, HBS- “Discrimination in Online Marketplaces” Recommended Reading: • Fixing Discrimination in Online Marketplaces by Ray Fisman and Michael Luca • Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment by Benjamin Edelman, Michael Luca, and Dan Svirsky Discussant: Ayelet Israeli, HBS (slides)

08:50 AM	**Algorithm Recommendations** | WCC 2004 - Harvard Law School

Megan Stevenson, George Mason- "Algorithmic Risk Assessment in the Hands of Humans"

Recommended Reading:

- Algorithmic Risk Assessment in the Hands of Humans by Megan Stevenson and Jennifer Doleac

Michael D. Ekstrand, Boise State- "Recommendations, Decisions, Feedback Loops, and Maybe Saving the Planet"

Recommended Reading:

- Effective User Interface Designs to Increase Energy-efficient Behavior in a Rasch-based Energy Recommender System by Alain Starke, Martijn Willemsen, and Chris Snijders
- Behaviorism in Not Enough by Michael D. Ekstrand and Martijn C. Willemsen

Katrina Ligett, The Hebrew University of Jerusalem- "Humans and algorithms, deciding together"

Discussant: Aaron Roth, University of Pennsylvania (slides)

01:00 PM	**Equality of Opportunity** | WCC 2004 - Harvard Law School John E. Roemer, Yale University – "Equalizing Opportunities through policy: A primer." Recommended Reading: • Equality of Opportunity: Theory and Measurement by John E. Roemer and Alain Trannoy • Equity in Health Care Delivery: Some Thoughts and an Example by John Roemer Rediet Abebe, Cornell University – "Mechanism Design for Social Good" Berk Ustun, Harvard University – "Actionable Recourse in Linear Classification" Recommended Reading: • Actionable Recourse in Linear Classification by Berk Ustun, Alexander Spangher, and Yang Liu Discussant: Rakesh Vohra, University of Pennsylvania (slides)

https://cra.org/ccc/wp-content/uploads/sites/2/2019/01/Algorithmic-and-Economic-Perspectives-on-Fairness.pdf

Report

Background Context

- Algorithmic systems have been used to inform consequential decisions for at least a century. Recidivism prediction dates back to the 1920s. Automated credit scoring dates began in the middle of the last century.

So what is new here?

- **Scale for one**
  - Algorithms are being implemented to scale up the number of instances a human decision maker can handle. Errors that once might have been idiosyncratic become systematic.
- **Ubiquity, is also novel**
  - Success in one context begets usage in other domains. Credit scores, for example, are used in contexts well beyond what their inventors imagined.
- **Accountability must be considered**
  - Who is responsible for an algorithm’s predictions? How might one appeal against an algorithm? How does one ask an algorithm to consider additional information beyond what its designers fixed upon?

Four Framing Remarks

- **One: The equity principle for evaluating outcomes**
- *Circumstances*, factors beyond an individual's control, such as race, height, and social origin
- *Effort variables*, factors for which individuals are assumed to be responsible.
- **Principle: Inequalities due to circumstances holding other factors fixed are viewed as unacceptable and therefore justify interventions**.

Four Framing Remarks

- **Two**: **Taste-based vs Statistical discrimination**
- *Taste-based*: discriminates against an otherwise qualified agent as a matter of taste alone
- *Statistical*: unconcerned with demographics *per se*, but understands that demographics are correlated with fitness for task

Four Framing Remarks

- **Two**: **Taste-based vs Statistical discrimination**
- *Taste-based*: discriminates against an otherwise qualified agent as a matter of taste alone
- *Statistical*: unconcerned with demographics *per se*, but understands that demographics are correlated with fitness for task
- Becker (1957): **taste-based discrimination is attenuated by competition** between decision makers with heterogeneity in taste.
- Policies to reduce statistical discrimination are less well understood.

Four Framing Remarks

- **Three: Emergence of Fair machine learning research**
- Goal is to ensure that decisions guided by algorithms are equitable.
- Over the last several years, myriad formal definitions of fairness have been proposed and studied.

Four Framing Remarks

- **Four: Mitigating data biases**
- Statistical ML relies on training data, which **implicitly encodes the choices of algorithm designers and other decision makers**.
- Can be a dearth of **representative** training data across subgroups
- **Target of prediction may be a poor** — and potentially biased — proxy of underlying act
- **Amplification**: When training data are the product of ongoing algorithmic decisions, feedback loops

Report Structure

1. Overview
2. Decision Making And Algorithms
3. Assessing Outcomes
4. Regulation and Monitoring
5. Educational and Workforce Implications
6. Algorithms Research
7. Broader Considerations

Decision Making And Algorithms

- “At present, the technical literature focuses on ‘fairness’ at the algorithmic level. The algorithm’s output, however, ***is but one among many inputs to a human decision maker***. Therefore, unless the decision maker strictly follows the recommendation of the algorithm, any fairness requirements satisfied by the algorithm’s output need not be satisfied by the actual decisions.”

Assessing Outcomes (1 of 2)

- “[because of feedback loops] in addition to good-faith guardrails based on expected effects, one *should also monitor and evaluate outcomes*. Thus, providing *ex ante* predictions is no less important than ***ex post*** **evaluations for situations with feedback loops**.”

Assessing Outcomes (2 of 2)

- "… a fundamental tension between attractive fairness properties… Someone's notion of fairness will be violated and tradeoffs need to be made... **These results do not negate the need for improved algorithms**. On the contrary, they underscore the need for informed discussion about fairness criteria and algorithmic approaches, tailored to a given domain.

Assessing Outcomes (2 of 2)

- "… a fundamental tension between attractive fairness properties… Someone's notion of fairness will be violated and tradeoffs need to be made... These results do not negate the need for improved algorithms. On the contrary, they underscore the need for informed discussion about fairness criteria and algorithmic approaches, tailored to a given domain. Also, these impossibility results are not about algorithms, *per se*. **Rather, they describe a feature of any decision process, including one that is executed entirely by humans**."

Regulation and Monitoring (1 of 2)

- “Effective regulation requires the ability to **observe the behavior of algorithmic systems**, including decentralized systems involving algorithms and people. … facilitates evaluation, improvement (including “de-biasing”), and auditing. … [but] transparency **can conflict with privacy considerations, hinder innovation, and otherwise change behavior**.

Regulation and Monitoring (2 of 2)

- “Effective regulation requires the ability to observe the behavior of algorithmic systems, including decentralized systems involving algorithms and people. … facilitates evaluation, improvement (including “de-biasing”), and auditing. … [but] transparency can conflict with privacy considerations, hinder innovation, and otherwise change behavior. Another challenge is that the disruption of traditional organizational forms by platforms (e.g., taxis, hotels, headhunting firms) has **dispersed decision making**. Who is responsible for ensuring compliance on these platforms, and how can this be achieved?”

Educational and Workforce Implications

- *“What should judges know about machine learning and statistics? What should software engineers learn about ethical implications of their technologies in various applications?* There are also implications for the interdisciplinarity of experts needed to guide this issue (e.g., in setting a research agenda). What is the relationship between domain and technical expertise in thinking about these issues? **How should domain expertise and technical expertise be organized**: within the same person or across several different experts?”

Algorithms Research

- "... a lot of work is happening around the various concrete definitions that have been proposed — even though **practitioners may find some or even much of this theoretical algorithmic work misguided**.

Algorithms Research

- "... a lot of work is happening around the various concrete definitions that have been proposed — even though practitioners may find some or even much of this theoretical algorithmic work misguided. How **to promote cross-field conversations** so that researchers with both domain (moral philosophy, economics, sociology, legal scholarship) and technical expertise can help others to find the right way to think about different properties, and even identify if there are dozens of properties whose desirability is not unanimously agreed upon?"

Broader Considerations

- “some discussion went to concerns about **academic credit and how the status quo may guide away from applied work**, noting also that the context of more applied work can be helpful in attracting more diverse students

Broader Considerations

- “some discussion went to concerns about academic credit and how the status quo may guide away from applied work, noting also that the context of more applied work can be helpful in attracting more diverse students
  … the research community **may ‘narrow frame’ the issues under consideration**. e.g., selecting from applicants those most qualified to perform a certain function is not the same as guaranteeing that the applicant pool includes those who might otherwise be too disadvantaged to compete.”

Visioning workshop: Algorithmic and Economic Perspectives on Fairness

Co-chairs: David Parkes (Harvard), Rakesh Vohra (Penn)

CCC Fairness and Accountability Task force: Liz Bradley, Sampath Kannan, Ronitt Rubinfeld, David Parkes, Suresh Venkatasubramanian