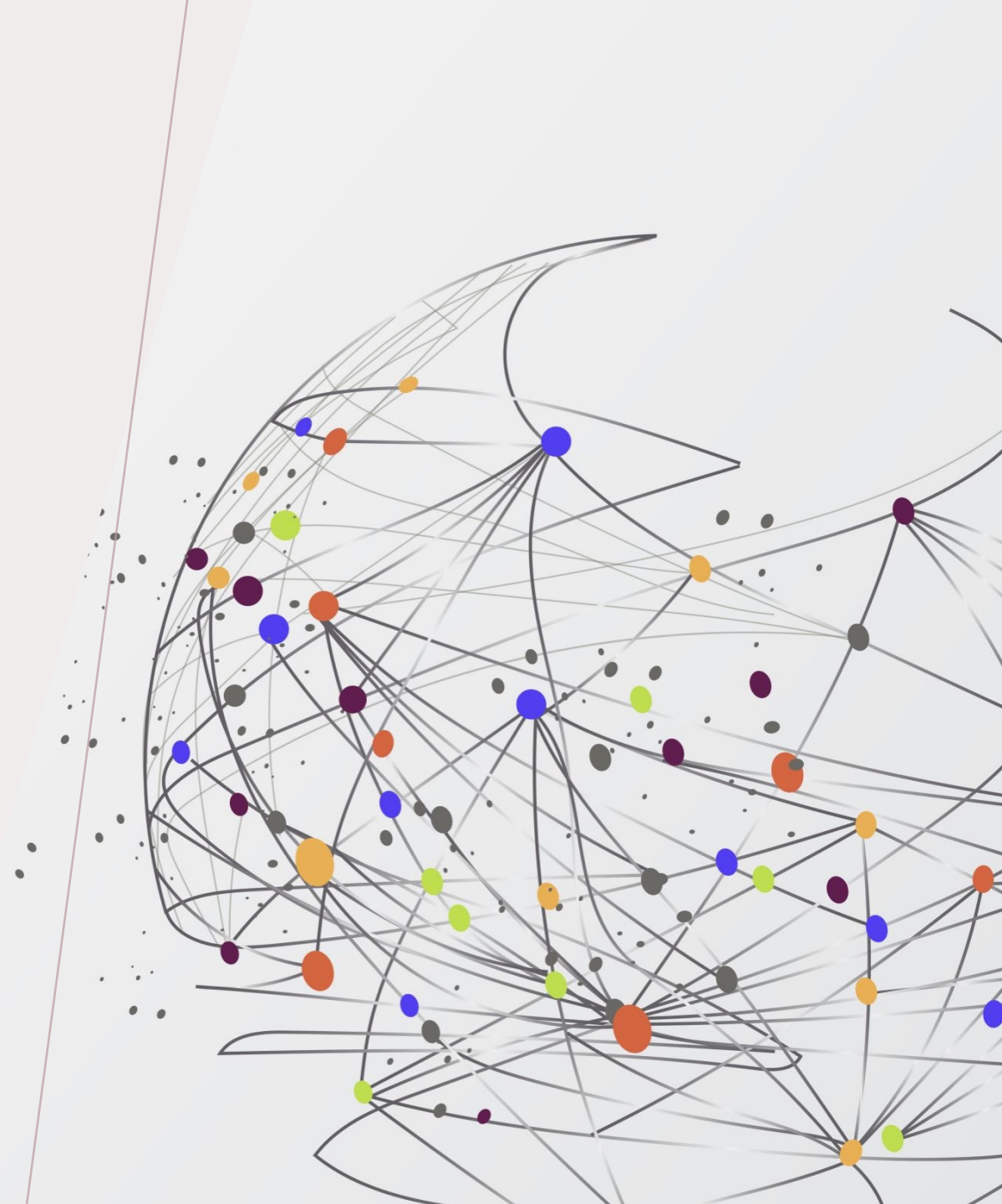


# The Possibilities of Building Technical Measures on Workplace Automated Decision- making Systems: **Taking 'Affirmative Action' Towards Algorithmic Discrimination**

**LUNA XIAOLU LI**

**J.D., Osgoode Hall Law School**

*Received research fund from Professor  
Valerio De Stefano (Canada Research  
Chair in Innovation, Law and Society)*



# Amazon built an AI tool to hire people but had to shut it down because it was discriminating against women

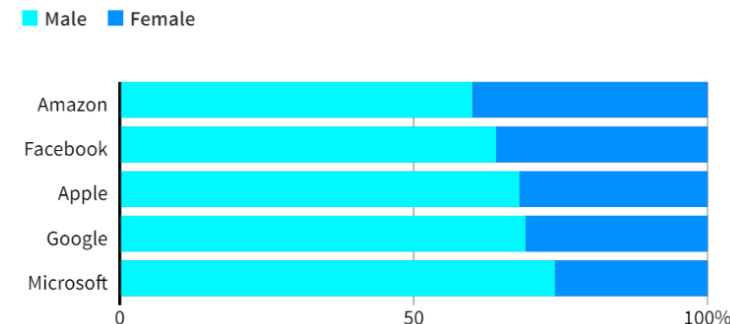
Isobel Asher Hamilton Oct 10, 2018, 5:47 AM



## Dominated by men

Top U.S. tech companies have yet to close the gender gap in hiring, a disparity most pronounced among technical staff such as software developers where men far outnumber women. Amazon's experimental recruiting engine followed the same pattern, learning to penalize resumes including the word "women's" until the company discovered the problem.

### GLOBAL HEADCOUNT



### EMPLOYEES IN TECHNICAL ROLES



Note: Amazon does not disclose the gender breakdown of its technical workforce. Source: Latest data available from the companies, since 2017.

## RESEARCH QUESTIONS:

1. Why is ADM systems popular? What are the problems?

2. Technical Measures on Algorithmic Discrimination?

3. Affirmative Action vs Racial Quota vs Racial Colour Blindness



# OUTLINE

## Background

- Definition of Algorithmic Discrimination & ADM Systems
- ADM's Popularity & Problems
  - **Discrepancy:** Panacea for Inefficiencies & Labour Shortages v.s. Black Box for Algorithmic Discrimination & Unfairness

## Reasons and Measures for Algorithmic Discrimination

- Direct & proxy effects
- Factors Leading to Algorithmic Discrimination
- Economical and Statistical Measures

## Legal Framework

- Legislation on Artificial Intelligence and Algorithmic Discrimination
  - Argue for a “Minorities-As-Whites” Model
  - Argue for Applying Affirmative Action to Anti-Algorithmic Discrimination
- Regulations in the Employment Context
  - Employers' Duties
  - Union's Involvements
  - Workers' Rights

## Conclusion



# ADM SYSTEMS' IMPACTS ON EMPLOYMENT FAIRNESS

## Pros & Cons on Employment Fairness

|   | Pros   | Cons  |
|---|--|---|
| Organizational Fairness v.s. Individual-Perceived Fairness      | <ul style="list-style-type: none"> <li>- Easier to discover hidden talented employees in organizations.</li> <li>- Able to screen a larger number of applicants automatically.</li> </ul>  | <ul style="list-style-type: none"> <li>- Individuals' perceived fairness of ADM is lower than human-conducted procedures in recruitment and HR development.</li> </ul>  |
| Consistency Among Individual Cases v.s. Systemic Discrimination | <ul style="list-style-type: none"> <li>- Firms seek to diminish the human biases by applying ADM, thereby increasing the objectivity, consistency, and employment fairness.</li> </ul>   | <ul style="list-style-type: none"> <li>- Algorithmic discrimination if ADM system is trained on inaccurate, biased, or unrepresentative input data.</li> <li>- Systemic discrimination even if prediction is accurate and consistent</li> </ul>   |
| Human Bias v.s. Algorithmic Bias                                | <ul style="list-style-type: none"> <li>- Possibility to be audited by examining the underlying code, or manipulating inputs and examining outcome differences.</li> <li>- Empirical evidence shows the potential to reduce human bias by ADM (though it is hard to completely remove algorithmic discrimination).</li> </ul> | <ul style="list-style-type: none"> <li>- Disclosure issue - trade secrets.</li> <li>- Legislation does not give a clear guidance on how to audit and monitor algorithms, or how to measure algorithmic discrimination. <ul style="list-style-type: none"> <li>❖ ADM developers and users may be motivated to simply remove gender/race data, leaving proxy effects that are hard to be detected.</li> </ul> </li> </ul> |

# ***IMPACTS ON EMPLOYMENT FAIRNESS***

## ***Reasons of Algorithmic Discrimination***

### **Reasons**

- **Unrepresentative Training Samples**
- **Mislabeling**
- **Proxy Effects (De Stefano 2018)**

***IMPACTS  
ON  
EMPLOYMENT  
FAIRNESS***

***Measures  
of  
Human Discrimination  
&  
Algorithmic Discrimination***

**Measures**

- *Prima Facie* Discrimination Legal Test (Canada):
  - *Human Rights Code*-protected ground
  - Adverse impact
  - Connection – high burden of proof for individual (esp. facing systemic discrimination)

## Measures – Algorithmic Discrimination

- Empirical evidence: measures exist for training samples & mislabeling → proxy effects can be mitigated.
- Key: ADM statistical model contains:
  - ✓ Part 1 - Training/Machine Learning
  - ✓ Part 2 - Prediction - on an individual basis
- Solution:
  - ✓ **Training data** contains demographic sensitive variables, whereas **prediction** of each new candidate is based only on non-demographic characteristics.
  - ✓ The non-demographic weights for this evaluation would come from the training process holding sensitive variables constant across all candidates.
  - ✓ **Machine version of affirmative action**



# Affirmative Action

- Ongoing lawsuit (2014-present)

*Students for Fair Admissions v President and Fellows of Harvard*







Proponents for affirmative action in higher education rally in front of the US supreme court on Monday. Photograph: Chip Somodevilla/Getty Images

# Supreme Court Seems Ready to Throw Out Race-Based College Admissions

The court's conservative majority was wary of plans at Harvard and the University of North Carolina that take account of race to foster educational diversity.



Activists speaking in support of affirmative action outside the Supreme Court after the justices heard arguments on Monday. Shuran Huang for The New York Times

**Update: Hearing before the Supreme Court of the United States on October 31, 2022.**

## The Minorities-as-Whites Model

- Yang & Dobbie 2020
  - Uses only the predictive power from inputs among White individuals (based on “Blinder-Oaxaca decomposition”)
- The algorithm treats minority individuals the same way it treats White individuals

***LEGAL  
FRAMEWORK***

***Regulation on  
Algorithm***

## Measures – Algorithmic Discrimination

- Empirical evidence: measures exist for training samples & mislabeling → proxy effects can be mitigated.
- Key: ADM statistical model contains:
  - ✓ Part 1 - Training/Machine Learning
  - ✓ Part 2 - Prediction - on an individual basis
- Solution:
  - ✓ **Training data** contains demographic sensitive variables, whereas **prediction** of each new candidate is based only on non-demographic characteristics.
  - ✓ The non-demographic weights for this evaluation would come from the training process holding sensitive variables constant across all candidates.
  - ✓ **Machine version of affirmative action**

## The Minorities-as-Whites Model

- Yang & Dobbie 2020
  - Uses only the predictive power from inputs among White individuals (based on “Blinder-Oaxaca decomposition”)
- The algorithm treats minority individuals the same way it treats White individuals

***LEGAL  
FRAMEWORK***

***Regulation on  
Algorithm***

## Employers' Duties

- Human Review;
- Internal & External Auditing (Ajunwa 2021);
- Workplace ADM Policy & Training

## Unions' Involvements

- Collective Bargaining

## Workers' Rights

- Data Access & Explanation
- To Contest & Negotiate

***LEGAL  
FRAMEWORK***

***Regulation on  
Employment /  
Labour Relations***



**Investigating AI can be one way  
to learn humanities.....**

**THANK YOU!**

**Luna Xiaolu Li**

