ALPEREN DEMİRCİ & ARDA DENİZ AYYILDIZ

ETHICS AND DISCRIMINATION IN AI BASED RECRUITMENT

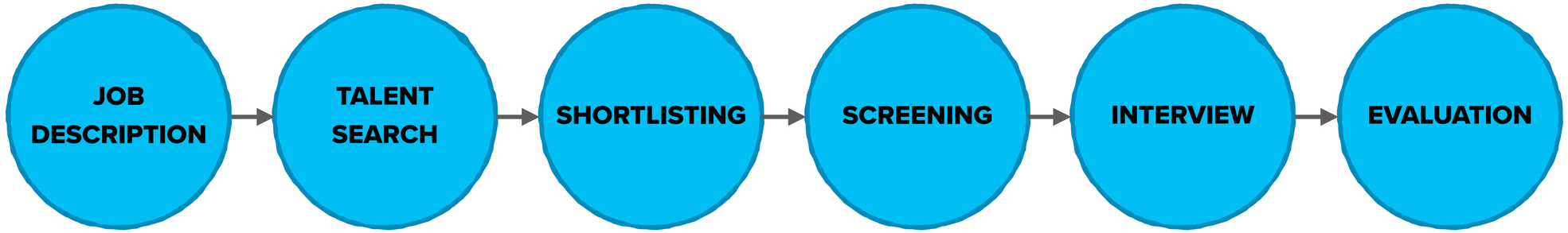
Did you really fail that interview just because you are not “good enough”?

WHAT WILL WE COVER

1. **Recruitment Processes and Discrimination Theory**
   - Understanding the terminology and identifying the problem.
2. **Benefits of AI-Based Recruitment**
   - Exploring the advantages of a smart, innovative approach.
3. **Discrimination in AI-Based Recruitment**
   - Humans are biased—so can AI be!
4. **Types of Discrimination in AI-Based Recruitment**
   - Identifying different forms of bias in AI systems.
5. **Real-Life Examples of AI Discrimination**
   - Case studies highlighting AI bias.
6. **Eliminating Discrimination in AI-Based Recruitment**
   - Strategies for creating fair AI recruitment processes.
7. **Legal Perspectives**
   - The regulatory landscape and legal considerations.
8. **Outline**
9. **Solutions**
   - Practical steps to address and prevent AI discrimination.
10. **References**

RECRUITMENT

- Difficult, expensive task to do. Can we automate this?
- Classic a decision problem with given attributes.
- **Recruitment Process**



- AI has affected nearly all industries, including HRM.
- Since recruitment is a part of HRM, companies start to develop AI based systems that will complete/help on some tasks of recruitment for recruiters.

DISCRIMINATION

DEFINITION

- "Discrimination in the labor market is defined by the ILO's Convention 111, which encompasses any **unfavorable** treatment based on **race, ethnicity, color, and gender** that **undermines** employment equality."

TYPES OF DISCRIMINATION

- **Competitive Market Theory:** This view attributes discrimination primarily to **personal prejudice**.
- **Monopoly Model of Discrimination:** It posits that **monopolistic power** leads to discriminatory behavior
- **Statistical Theory of Discrimination:** It suggests that **nonobjective** variables, such as inadequate information, contribute to biased outcomes.

- Since we will talk about AI, which is a bunch on algorithms that runs with data, the **Statistical Theory of Discrimination** is important for us.

BENEFITS OF AI BASED RECRUITMENT

- **Recruitment Quality**
  - Large companies believe unconscious bias affects recruitment quality, and AI can help by automating the process to ensure **better candidate matching**.
  - This technology improves job descriptions, targets advertisements, and automates resume screening, making hiring **more efficient** and **unbiased**(?).
- **Recruitment Efficiency**
  - HR departments receive many candidates for each position, making traditional human-based screening **expensive** and **discouraging**.
  - AI can speed up hiring, enhance the candidate experience, and reduce costs by quickly delivering job information, screening out **uninterested applicants**, and automating **resume evaluation** and **candidate classification**.
- **Transactional Workload**
  - AI in recruitment is seen as a "new era in human resources" because it replaces **routine tasks** performed by human recruiters, transforming traditional practices.
  - AI reduces administrative tasks, letting recruiters focus on **strategic affairs** like recruiting, screening, and interviewing.

DISCRIMINATION IN AI BASED RECRUITMENT

- **Datasets**
  - Datasets are the foundation of ML, and if they lack quantity and quality, bias in algorithmic decisions is inevitable.
  - Skewed datasets often reflect imbalances in gender and race, leading to biased hiring processes.
  - Insufficient representation in data can exclude underrepresented groups and perpetuate bias towards candidates resembling existing employees.
- **Impact of Unstructured Data**
  - Most datasets consist of unstructured data from observational measures, which can result in misreporting.
  - Algorithms favor the represented groups and operate less effectively for underrepresented groups.
  - Existing social biases in raw data lead to the "bias in and bias out" phenomenon, where historical inequalities are projected and amplified.
- **Designer Bias**
  - Bias can be introduced in model construction if computer scientists are unaware of social issues.
  - Engineers play a crucial role in setting goals, selecting models, and determining data characteristics, which can introduce bias.
  - Personal biases of engineers can influence the algorithm's decisions, as seen in the Amazon hiring(2014) case where gender was considered a crucial criterion.

WHICH DISCRIMINATIONS EXIST?

- **Gender**
  - Gender stereotypes have affected NLP and ML, leading to biased job search results and making minority-gender jobs seem less professional.
  - For example, Amazon's 2014 ML-based hiring tool exhibited gender bias by downgrading applicants with keywords such as "female," leading to its withdrawal and redevelopment.
- **Race**
  - Microsoft's chatbot Tay learned to produce sexist and racist remarks by interacting with Twitter users, leading to its shutdown within hours.
  - Additionally, searches for names associated with Black individuals often triggered arrest record ads, while names associated with white individuals did not, reflecting subconscious biases in AI.
- **Skin Color**
  - In 2015, Google's photo app mislabeled a photo of two Black people as gorillas due to insufficient training on diverse skin tones.
  - Similarly, in 2017, an algorithm for a contactless soap dispenser failed to detect black and brown hands, highlighting issues of algorithmic bias.
- **Personality**
  - Algorithms assess word choice, tone shifts, and facial expressions to determine a candidate's personality and cultural fit.
  - For instance, they might correlate job tenure with "high creativity" or link curiosity to a likelihood of seeking other opportunities, using sentiment analysis to gauge emotional tone.
  - Another example is that, one resume screening tool identified being named Jared and playing high school lacrosse as correlated with being a successful employee.

DISCRIMINATION WITH REAL LIFE EXAMPLES

- **Open AI - ChatGPT 3.5 for Recruitment Task**
  - A Bloomberg investigation found that OpenAI's GPT-3.5 showed racial and gender biases in hiring, disadvantaging Black women in technical roles.
  - The study revealed that the AI inconsistently ranked candidates based on demographics(names).
  - OpenAI stated that clients often customize the software to reduce bias, but the findings raise concerns about discrimination in AI hiring tools.
- **Money Bank - California.AI lawsuit**
  - Money Bank used an AI tool, GetBestTalent from California.AI, for recruitment, which led to claims of discrimination.
  - Alice (female), Frank (Black), and James (61 years old) all felt they were unfairly rejected due to biases in the AI system.
  - The AI tool had been advertised as bias-free, yet these rejections highlighted significant issues, prompting legal challenges based on sex, race, and age discrimination.
- **COMPAS (Correctional Offender Management Profiling for Alternative Sanctions)**
  - COMPAS algorithm is used in US court systems to predict the likelihood that a defendant would become a recidivist.
  - Due to the data that was used, the model that was chosen, and the process of creating the algorithm overall, the model predicted twice as many false positives for recidivism for black offenders (45%) than white offenders (23%).

HOW TO ELIMINATE DISCRIMINATION ON ABR?

- **Constructing a more unbiased dataset**
  - Unfair datasets cause bias, so fixing them is crucial. Microsoft did this for its Face API, reducing errors for darker skin tones by balancing skin color, age, and gender.
  - Combining "small data" with "big data" boosts accuracy by focusing on precision. MIT researchers showed that an AI system can reduce bias by re-sampling data, improving fairness in facial recognition.
- **Enhancing Algorithmic Transparency**
  - Transparency in AI helps solve the "algorithmic black box" problem, making bias easier to fix. Engineers often struggle to understand AI outcomes, leading to unintentional discrimination.
  - Tools like Blendoor reduce unconscious bias by removing identifiable info from candidate profiles. Deep Xplore finds and corrects biased algorithms in neural networks through discrepancy testing.
- **Internal Ethics Governance**
  - Tech companies are self-regulating with AI principles to address bias. Microsoft formed an AI standards & Ethics committee, and Google introduced Model Cards to explain algorithms and their results.
  - Internal audits and diverse data collection help correct biases. For example, AI-HR audits review employee selection processes and assess AI reliability.
- **External Supervision**
  - Third-party certification and testing ensure AI transparency and accountability. Public organizations like the Algorithm Justice League improve algorithm accuracy in minority facial identification.
  - Non-discrimination laws like the EU's GDPR protect against biased algorithms. The GDPR allows individuals to request explanations for algorithmic decisions and mandates Data Protection Impact Assessments.

LEGAL POINT OF VIEW

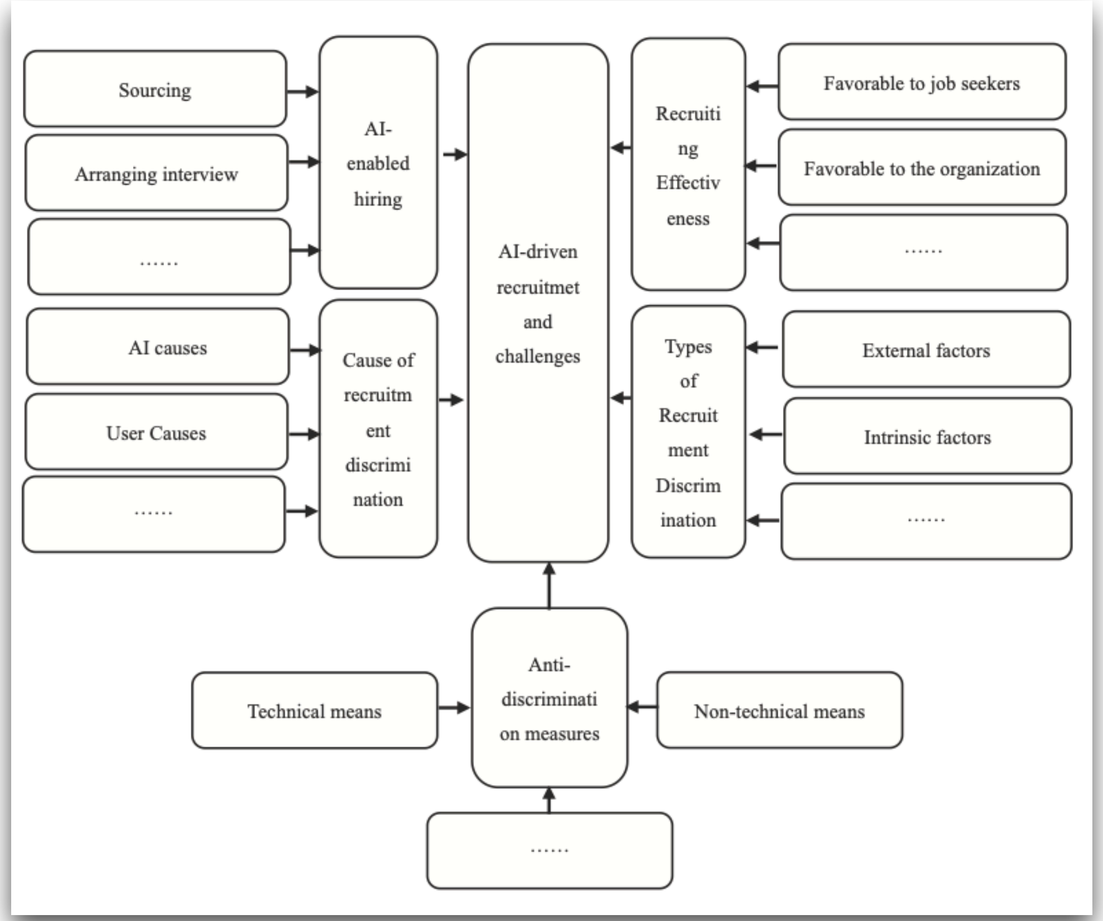
- **For Turkiye**
  - Turkiye does not have a specific law about algorithmic bias, however it has TİHEK 6071 which is against all types of discrimination.
  - The Turkish Human Rights and Equality Institution Law No. 6701 aims to prevent discrimination and promote human rights based on human dignity.
  - It establishes the Human Rights and Equality Institution of Turkey, defining its organization, duties, and authorities to ensure equal treatment, prevent discrimination, and combat torture and ill-treatment effectively.
- **For USA**
  - NYC 144, stipulates that employers who use an automated employment decision tool or “AEDT” (automated hiring system) to assist with employment decisions must verify whether such tools are acting biased.
  - NYC 144 shortly is a compliance law for employers to make “AEDT” processes more clear and transparent.
  - Some lawmakers want to make employers using AI tools follow stricter non-discrimination rules.
  - The American Data Privacy and Protection Act bans using data in ways that discriminate based on race, color, religion, national origin, sex, or disability.
  - Some states also want to ban very risky AI tools.

WHAT DO INTERVIEWEES THINK ?

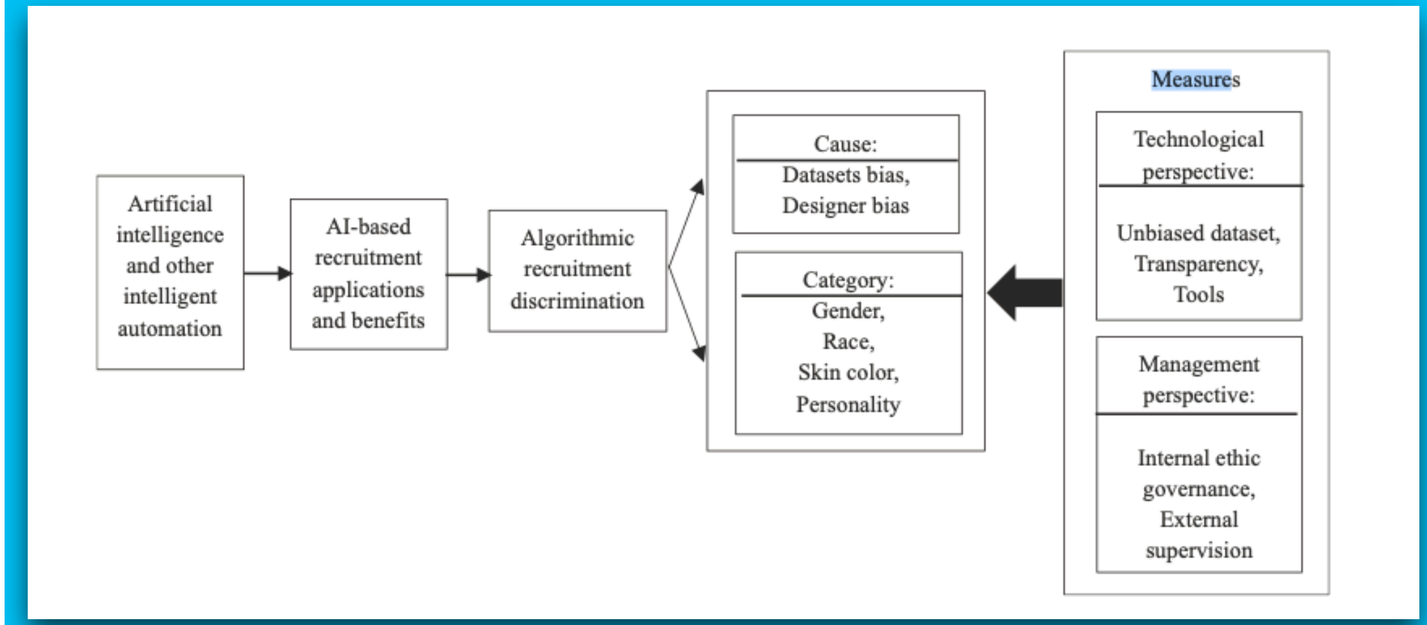
- **Experiment:**
  - Bunch of interviewees are interviewed using a qualitative analysis software (Nvivio 12.0 Plus).
  - This software had some features which is powered by AI, like transcription and coding of the interview recordings or filtering the irrelevant applicants based on their skills/jobs etc.
  - The interviewees are interviewed for 30 minutes with different meetings(face to face, online, telephone).
  - At the end, researchers asked what do job seekers feel about usage of AI in recruitment processes.
- **Results:**
  - In the end, our job seekers concluded that there are 5 key points for this topic:
    1. Time
    2. Human Bias
    3. Unfamiliarity
    4. Algorithmic Bias
    5. Fighting with Bias

OUTLINE OF OUR COVERAGE

- Simple schema of what we have covered so far.



- Timeline for the ABR system and its' effects.



REFERENCES

- https://www.aclu.org/news/racial-justice/how-artificial-intelligence-might-prevent-you-from-getting-hired
- https://mashable.com/article/chatgpt-ai-racism-bias
- https://www.lewissilkin.com/en/insights/discrimination-and-bias-in-ai-recruitment-a-case-study
- https://mashable.com/article/openai-chatgpt-racial-bias-in-recruiting
- https://kplegal.com.tr/en/an-interesting-decision-on-artificial-intelligence-driven-recruitment-tools-the-bias-law-for-employment-algorithms-and-the-situation-in-turkey/
- https://www.logically.ai/articles/5-examples-of-biased-ai
- https://datatron.com/real-life-examples-of-discriminating-artificial-intelligence/
- https://www.researchgate.net/publication/373948488_Ethics_and_discrimination_in_artificial_intelligence-enabled_recruitment_practices
- https://doi.org/10.1057/s41599-023-02079-x