Using AI to promote equity in the job market

Professor Roger A. Søraa

Principal Investigator (PI)

Just?Tech online lecture

20.May 2025

BIAS — Mitigating biases of AI in the labour market

NTNU — Norwegian University of Science and Technology

AI Fairness Cluster

Hi, I'm Roger A. Søraa

- Professor at the Department of Interdisciplinary Studies of Culture (KULT) in Science and Technology Studies
- My focus is on how humans and technology relate to each other
- In a work context – especially AI

Author of *AI for diversity* (2023) →
- AI has great potential, but can also lead to exclusion for different groups
- How the AI is designed and how data is selected and trained are not neutral factors, but can lead to biased outcomes
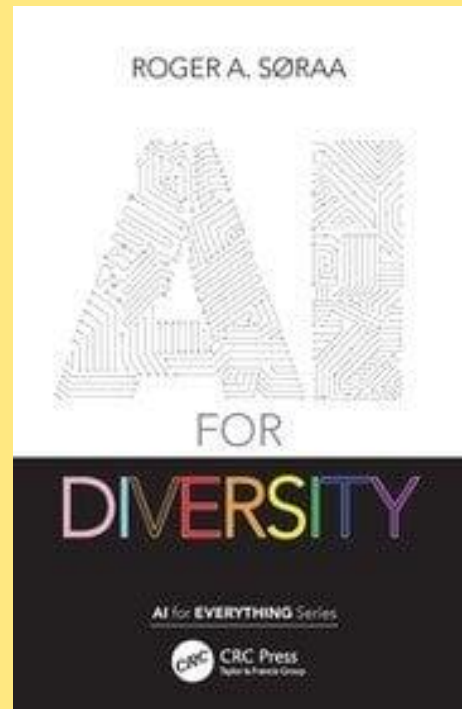
The BIAS project

- Funded by the European Union's *Horizon Europe* research program

- Under Pillar 2, Cluster 4

- Total project size: € 4.2 million

- Led by NTNU Humanities Faculty by PI Prof. in Science and Technology Studies (STS) Roger A. Søraa
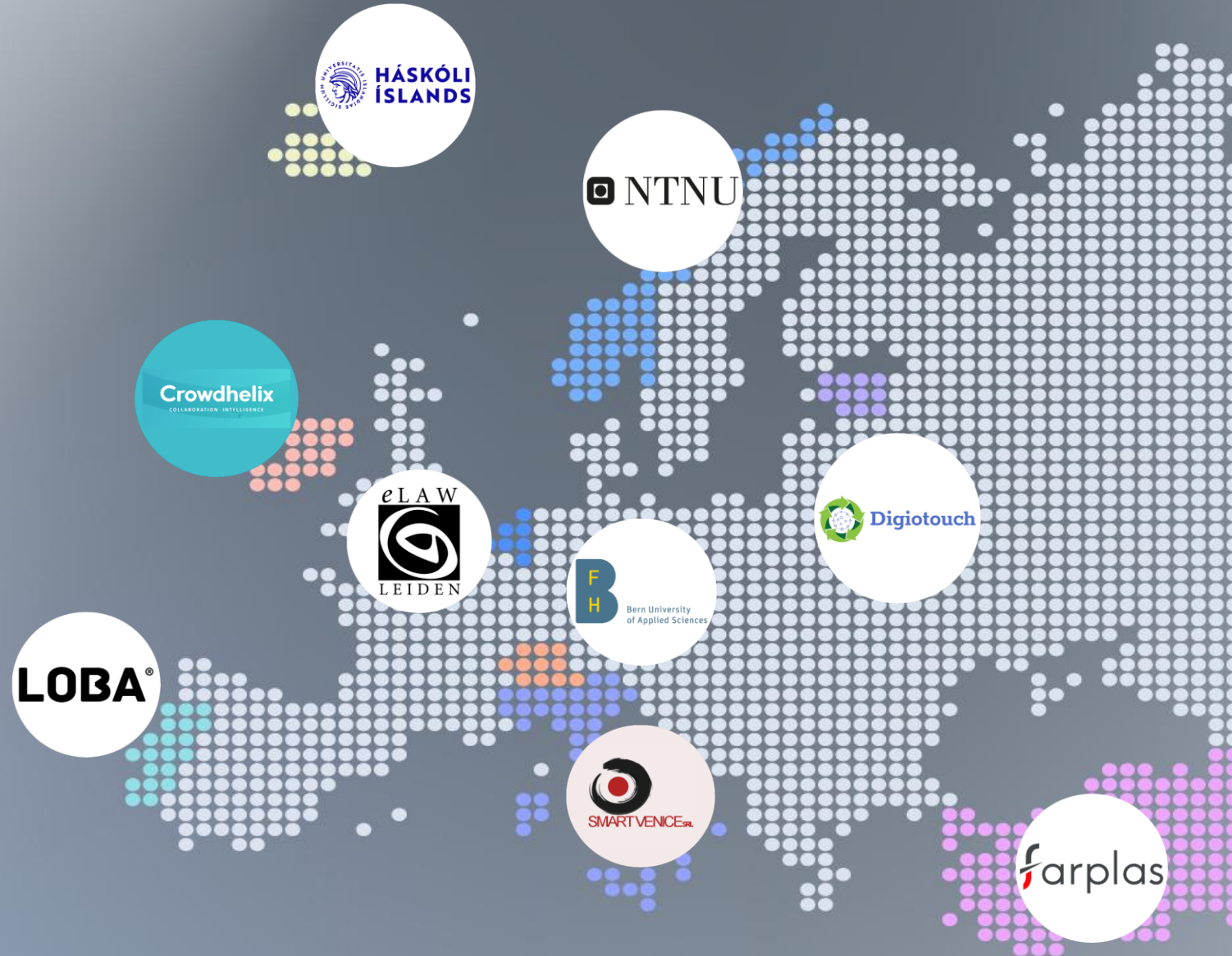
- From 11/2022 - 10/2026 (4 year)

BIAS
CONSORTIUM MAP

- ● **NORWAY** – NTNU
- ● **SWITZERLAND** – BFH
- ● **ICELAND** – HI
- ● **PORTUGAL** – LOBA
- ● **IRELAND** – CHX
- ● **ITALY** – SVEN
- ● **NETHERLANDS** – LEID
- ● **ESTONIA** – DIGI
- ● **TURKEY** – FARPL

BIAS Project mission

Research & empower the **AI** and **Human Resources Management** (HRM) community by addressing and mitigating AI biases.

Sociotechnical interdisciplinary knowledge on job seeking process

How to ensure:

Research show **humans** discriminate when hiring,
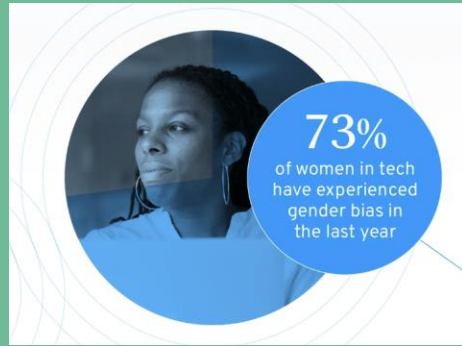e.g. based on:
-age, gender, race, class

BIAS

Images from: Creative Commons

Women in Tech Statistics:
73% Experience Gender Bias in the Workplace (https://spr.com/)

73% of women in tech have experienced gender bias in the last year

Nationality discrimination in Italy: 2018, Italians applying for a job received a much higher proportion of replies than immigrants. This discrimination **intersected with gender dimensions**, with Italian women receiving less replies than Italian man (9.2% callback rate for women compared with 14.5% for man) and **immigrant women** receiving the lowest number of replies (with **3.9% callback rate** for first generation immigrant Moroccan women).

A history of human discrimination at work

"Applications with non-Norwegian sounding names 25% lower chance to be called in for job interviews"

Det er fortsatt betydelig vanskeligere å få et jobbintervju hvis du har utenlandsk navn. Det viser en ny forskningsrapport.

Bank fined for canceling hiring pregnant applicant in Türkiye

BY DAILY SABAH WITH AA | ISTANBUL | OCT 10, 2023 - 11:12 AM GMT+3 |

A pregnant woman works on a laptop while sitting at her workplace in an office. (Shutterstock Photo)

"Bank in Turkey retracts job offer after finding out new hire was pregnant"

Age discrimination in Belgium: " the younger candidate within a pair of applicants had, on average, a 64.3% higher chance of being invited for a job interview than the older candidate, and a 39.7% higher chance of receiving some sort of positive reaction"

* Baert, S., Norga, J., Thuy, Y., & Van Hecke, M. (2016). Getting grey hairs in the labour market. An alternative experiment on age discrimination. Journal of Economic Psychology, 57, 86-101. // Busetta, Giovanni, Maria Gabriella Campolo, and Demetrio Panarello. "Immigrants and Italian Labor Market: Statistical or Taste-Based Discrimination?" Genus 74, no. 1 (February 22, 2018): 4. https://doi.org/10.1186/s41118-018-0030-1. [...]

Artificial Intelligence (AI)
is increasingly used for **recruitment**,
e.g. on:
-finding, screening and assessing candidates

BIAS

What does a text tell you about an applicant?

Whereas data in texts such as motivation letters in applications are anonymized to avoid bias, this is often not enough.

Some **patterns not explicitly visibile to the humans might be picked up by the AI and lead to a bias** in decision making upon such documents.

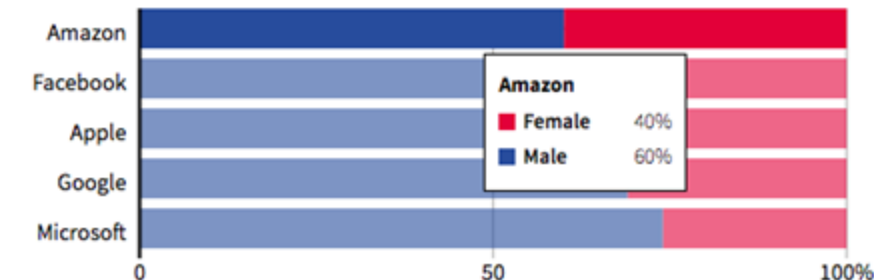This is in particular delicate for the matter of automated recruitment tools.
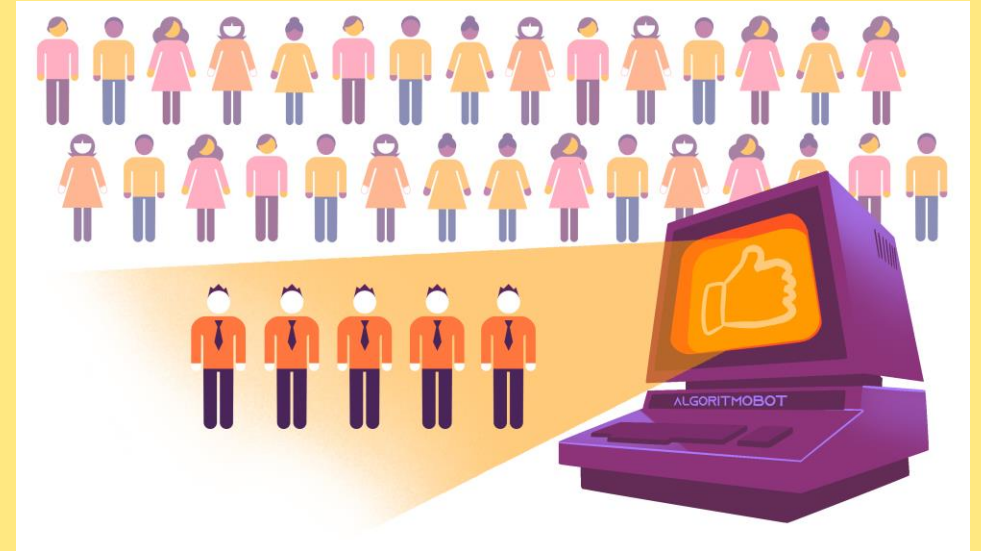
Created by: Nienke Bruijning

BIAS

AI bias for worker recruitment

- In 2015, Amazon's AI recruiting system showed bias against women.
  - Historical data from the last 10-years to train their AI model.
  - Male dominance across the tech industry
  - The AI "learned" that recruiting male candidates were preferable.
  - Tagged word as negative, e.g. "women's chess club captain."

- In 2019, bias in job adverts, where Facebook targeted ads intentionally target by gender, race, age, religion etc.
  - Women shown job adverts for nursing or secretarial work,
  - Men shown job ads for janitors and taxi drivers (in particular men from minority backgrounds). (Dilgemi. 2020)

GLOBAL HEADCOUNT
■ Male   ■ Female

Amazon
Facebook
Apple
Google
Microsoft

Amazon
■ Female   40%
■ Male   60%

0   50   100%

Source:
https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/

Other ways AI is used

- AI scans online platforms (e.g., LinkedIn) to identify potential candidates based on descriptions and past hires.

- Ranks resumes using natural language processing (NLP) based on experience, education, skills, and keywords.

- Analyse video interview based on: tone of voice, facial expressions, choice of word

AI in recruitment can reinforce existing biases, obscure decision-making processes, and reduce transparency—especially when algorithms rely on historical data or opaque criteria like facial expressions or keyword frequency. Therefore, it is important that we keep in mind the need for fairness, accountability, and explainability throughout the hiring process.

BIAS

Image from: Creative Commons

How can we:
 (1) make sure that AI does not **reproduce human** bias & discrimination, nor that
(2) AI **invents new** forms of discrimination for job application and workers?

Preliminary project results from the BIAS project

Survey with 4 000+ workers across Europe concerning perceptions and experiences of AI at work

Co-creation and capacity building to align technology development with HR needs and mapping country & language-specific indicators of bias in HR context

Co-creation complete; capacity building ongoing

200+ in-depth interviews with workers, HR managers, AI developers and 10 long-term in-depth site visits of companies about how AI is developed and deployed

Proof-of-concept NLP and CBR engine for identification and mitigation, utilizing real world & synthetic data, focusing on non-English languages

under development

AI systems reproduce human bias if we are not aware of:
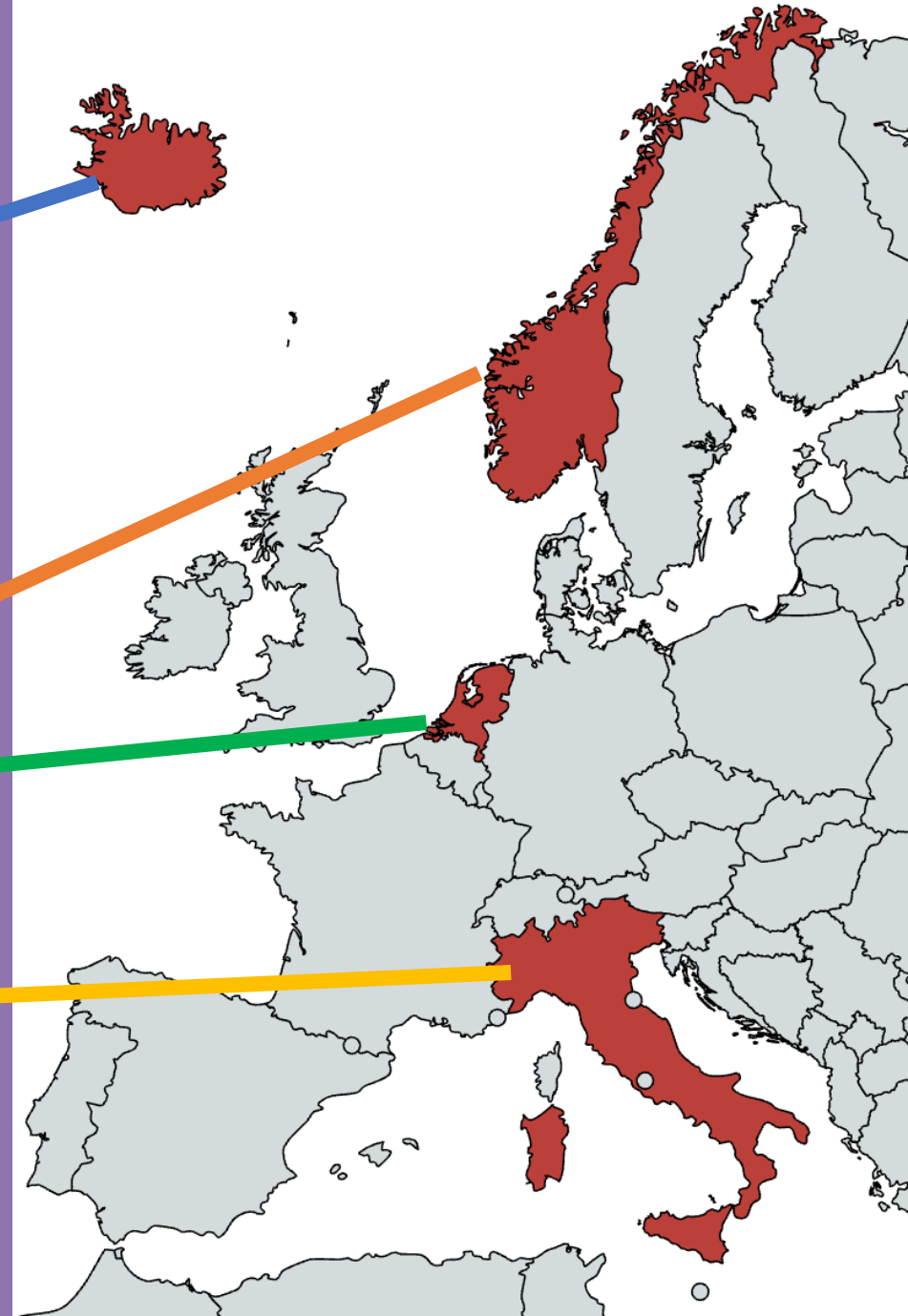
Fieldwork on HR practitioners in Italy

Focus: HR practitioners' attitudes towards AI for recruitment

- 3 months of fieldwork: 26 interviews, 2 focus groups, 64 total hours of observation – **58 total participants**

- 2 field sites: 1 head-hunting company using AI CV ranking & summary, 1 HR department of an NGO using AI to write candidate's synthetic evaluations.

Findings:

- HR practitioners are aware of the risks of human bias in recruitment but **trust their own sensibility** and try to mediate between job seekers' and company's interests.

- HR practitioners trust their "gut feeling" can be a sign of experience but can also **potentially lead to bias**, especially affinity bias.

- AI is spreading for **CV screening and ranking** – but people still believe interviews should be done by people

- HR practitioners are aware of the risk of AI bias, hence, they **oversee algorithms** as much as possible

- Expectations about the future role of AI in HR are intertwined with **expectations about work and the HR role**.

BIAS

Fieldwork on Chinese job seekers in Norway and the Netherlands

- Individual case studies following Chinese job seekers, key themes:
  - Struggles and barriers in job seeking as immigrants
  - AI leverage in job seeking process
  - Perceptions on fairness, diversity and inclusion
- Findings
  - **Language skills** and **working vis**a are two main challenges.
  - **National security regulations** and **geopolitical restrictions** are tricky to be categorized as discrimination.
  - Job seekers are in **a vulnerable state** both mentally and socially to fight with unfair treatment.
  - Whether AI is used by recruiting company is **speculation from job seekers**.
  - AI usage from job seekers is considered as **a resistance** toward AI usage from large companies.
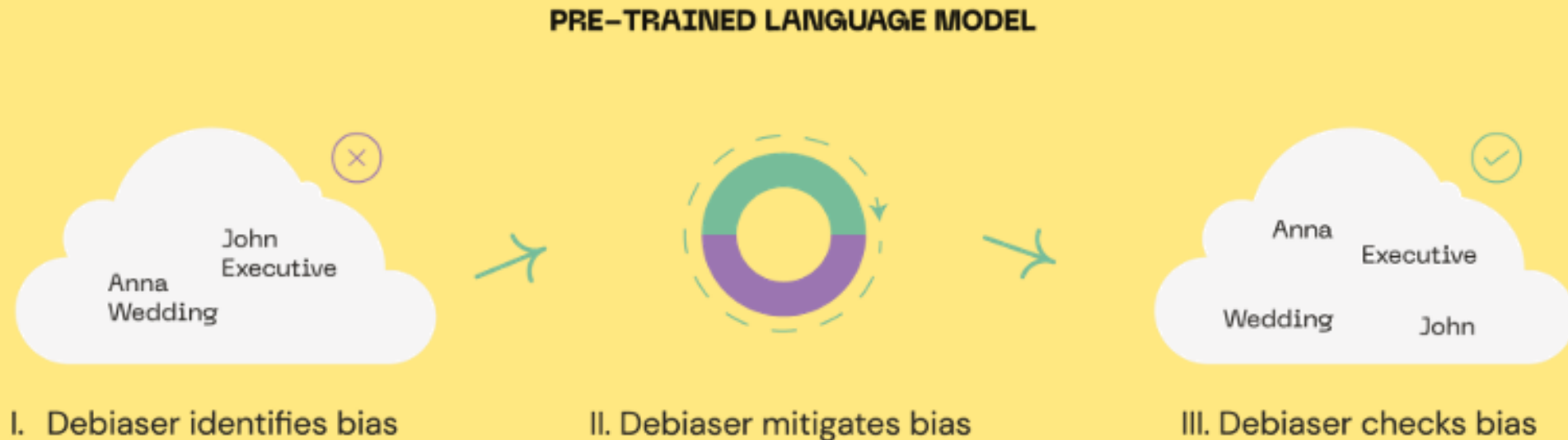
BIAS

Example: Human bias in hiring

- *"We don't see gender, race, class etc., but only meritocracy."*
- Similarity attraction bias (the candidate looks like oneself)
  - Halo effect, Horn effect,
  - (some others: Confirmation bias, Affect heuristics, Expectation anchor, Overconfidence bias, Beauty bias Conformity bias+)
- A majority blindspot, difficult to advance as "the other"
- What happens when computer systems are made to assess workers and candidates for jobs?

Image from: Creative Commons

BIAS

The Debiaser tool

The Debiaser provides a **toolkit to identify and mitigate the bias** in such language models, making them safer to use in your application. In combination with our **expertise** from research, the toolkit and application is adapted exactly to organizational needs, in terms of language or application use case.

PRE-TRAINED LANGUAGE MODEL

I. Debiaser identifies bias

II. Debiaser mitigates bias

III. Debiaser checks bias

BIAS

What is **fairness**?

When someone applies for a job, how can you make sure that the **right person** got the job – and ensure a **fair hiring** process?

BIAS

The problem of "fairness"

- The challenge of bias in automated recruiting starts with the definition of **what fairness means**.

- Different **disciplinary** understanding

- Different **socio-cultural** understandings

- We need to continue discussing what is fair:
  - For individual?
  - For the company?
  - For a minority group?
  - For society?

- We look at procedural fairness for individuals—that **similar candidates are treated in a similar way**—AND group fairness for **how AI treats minority groups**.
  - What similar means, is **dependent on individual businesses** and use cases.

Image from: Creative Commons
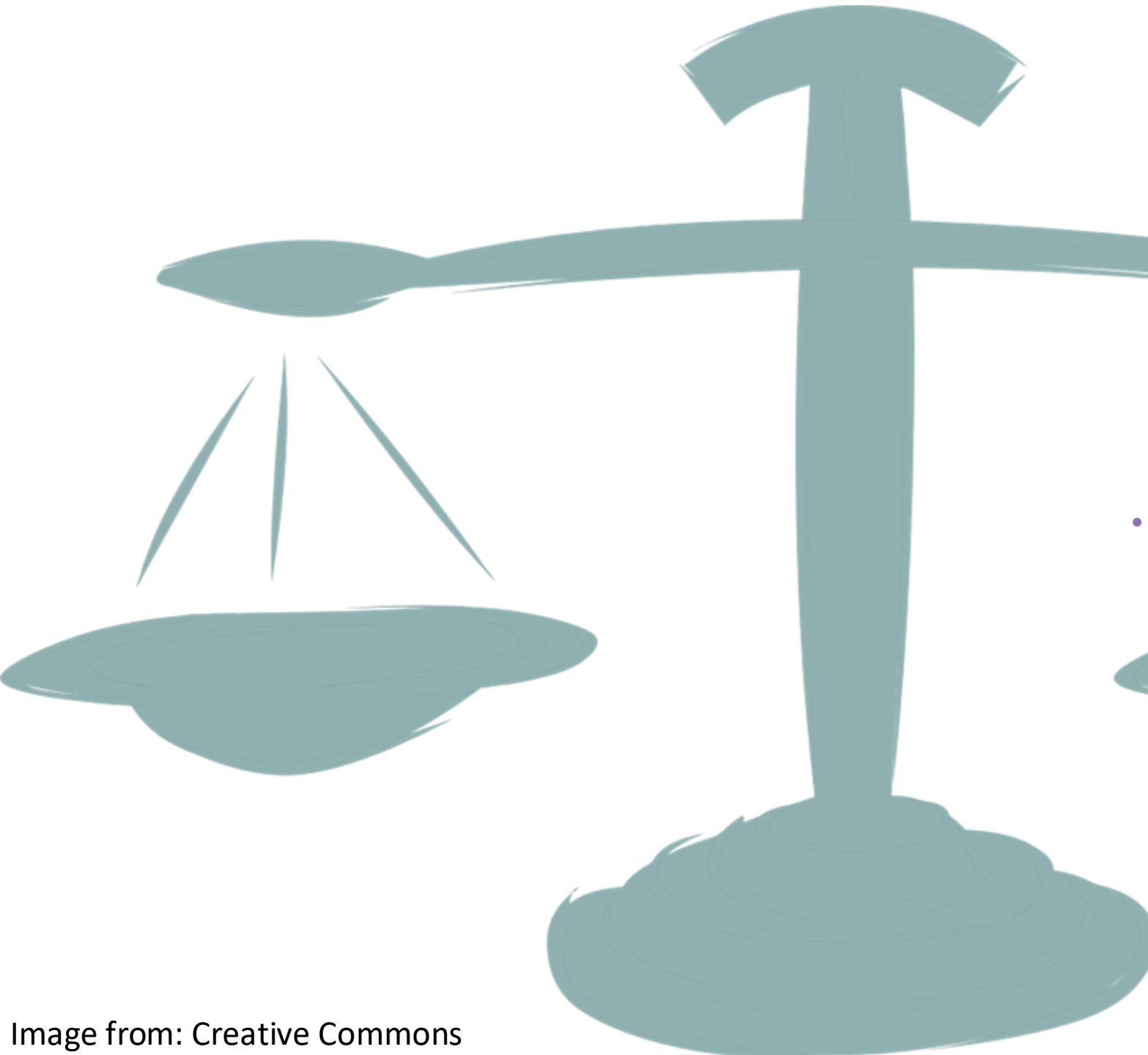
How can AI be used for good when recruiting?

- Traditional human hiring isn´t neutral. We all carry unconscious biases shaped by culture, background and experience

- By recognizing existing biases in recruitment, we can train AI to ignore irrelevant factors (e.g., gender, ethnicity) by applying consistent criteria

- Examples of AI built for diversity:
  - AI-based writing tool to craft more inclusive job ads
  - Standardize candidate assessments to reduce human bias
  - Anonymization of sensitive information from candidates
  - Offer alternative evaluation criteria that go beyond traditional human screening

- AI systems continue the patterns in their training data. Therefore, human supervision will always remain important

BIAS

Image from: Creative Commons

Life is "unfair" – but we can mitigate this!

Reality	Equality	Equity	Impartial	Inclusion
Some gets more than is needed, while others get less than is needed	The assumption is that everyone benefits from the same support	Everyone gets the support they need, which produces 'equity'	Causes of inequity addressed. Systemic barrier has been removed	Everyone is included and involved

Funded by the European Union

BIAS

Image from: Future institute of Australia

AI to advance the Sustainable Development Goals (SDGs) and tackle global challenges

- AI for Good was established in 2017 by the International Telecommunication Union (ITU) in partnership with over 40 UN Sister Agencies wit the aim of **identifying innovative AI applications, building skills and standards, and advancing partnerships to solve global challenges**.

- AI for **SDG 8: Decent Work and Economic Growth**

- AI can provide increased accessibility to work
  - Also in recruitment! For example, ensure more people are shown job adverstisements

- AI supporting decision-makers with data gathering, summarization and analysis

BIAS

Examples from BIAS research

AI to make recruitment fairer

- Co-creation workshop participants came up with the idea of using AI to draft feedback messages to all candidates, explaining why they have not been selected and what they could improve.
  - This can be very useful in contexts like Italy, where sending feedback messages to candidates is not a common practice and job seekers are often left wondering

- 2 cases of Italian recruiters becoming aware of their own biases thanks to the AI system:
  - A recruiter realized he had an unconscious bias against people who went to online universities, he was able to overcome this thanks to the AI CV ranking that place a candidate who went to a online university first
  - A recruiter realized her own biases when comparing the input she gave to AI with her colleagues in how to evaluate interview transcripts. She realized she evaluated behaviours differently than her colleagues, preferring a certain managerial style over another due to her personal preference.

BIAS

Summary

- Hiring is often discriminatory when done by humans

- We need to ensure that AI does not *also* discriminate (and mitigate human discrimination)

- Ensuring fair hiring practices is difficult (both for humans and machines)

- Ethical AI is good in theory – but how to ensure practical implementation?

BIAS

BIAS

Mitigating biases
of AI in the
labour market

CONTACT

Prof. Roger Andre Søraa

roger.soraa@ntnu.no

@rogerSoraa

thank you

Join our national labs
& learn more →

European
Commission