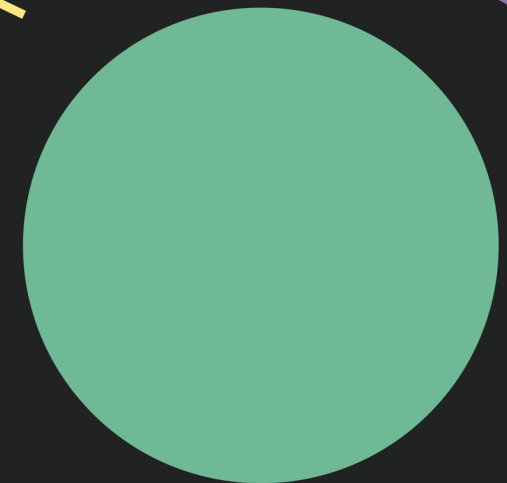
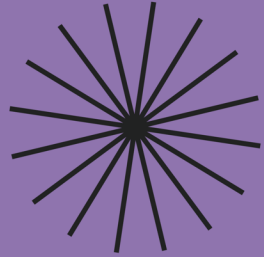


BIAS

Mitigating biases
of AI in the
labour market





Welcome to BIAS

A European funded project that will empower the Artificial Intelligence (AI) and Human Resources Management (HRM) communities by addressing and mitigating algorithmic biases.



Funded by
the European Union

BIAS



Why AI and HRM?



Funded by
the European Union

BIAS



In the modern world, AI is increasingly being used in the job market to engage and train employees, monitor infractions, and even make hiring decisions (according to a recent study by the consultancy Sage, approximately 24% of companies use AI for hiring purposes).

One type of AI tool is Natural Language Processing (NLP), a technology that helps computers understand and work with human language. It involves teaching computers to read, interpret, and respond to words and sentences, just like we do as humans.



EXAMPLES

AI in HRM can be used for:

- Generating job postings;
- Processing large amounts of job applications by sorting or ranking them;
- Automatically extracting information from cover letters;
- Manage employees by generating shifts, monitoring schedules, working hours, efficiency or other activities.

The information gathered through AI can either result in an automated decision-making process about who to invite for an interview, who to hire or who to promote or to inform such a decision.



But what if these AI-based systems are biased? What if the algorithms reflect societal stereotypes, leading to unfair decisions?

Bias in the recruitment process can take various forms, such as gender, racial, age, educational, socioeconomic, unconscious, confirmation, and halo/horn bias. These biases can lead to unfair treatment, exclusion of qualified candidates, and inequality in the workplace.



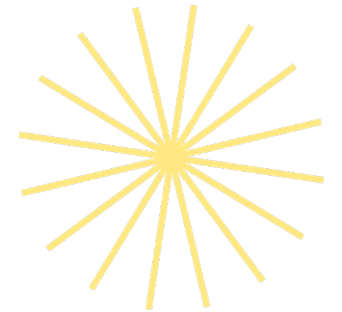
EXAMPLES

One example of this is the Amazon biased hiring system

In 2014 the company Amazon developed an AI-based tool for hiring and started testing it for their operations. However, the company decided to withdraw the tool in 2015 once they realized it had a negative bias against women.

Even if data about the gender of applicants was hidden, the algorithm picked-upon other markers which resulted in gender-based discrimination. The AI penalized resumes including the word "women's" and downgraded graduates from two all-women's colleges of the United States.

This is a clear example of how the introduction of AI for HRM, even if done with the best of intentions, could lead to negative effects.





Bias in AI for HRM can be attributed to various factors:

- AI trained on uncorrected historical data tends to replicate and magnify real-world biases. For instance, if engineering positions at Amazon have predominantly been held by males in the past, the algorithm may inadvertently incorporate this bias into its system.
- Even when sensitive data is concealed, AI can generate biased results by relying on proxy variables. In the case of Amazon, although the candidates' gender was concealed, the algorithm penalized applicants who had attended all-women's colleges.

The BIAS project aims to respond to these challenges by empowering the AI and HRM communities to address and mitigate algorithmic biases.



Main objectives of the project



Funded by
the European Union

BIAS



OBJECTIVES

01



Develop tools to identify and reduce bias in AI/NLP systems

03



Understand biases in recruitment to improve HRM practices

05



Reduce biases from the hiring process by creating more equal and fairer practices

02



Empower AI and HRM communities to reduce algorithmic bias

04

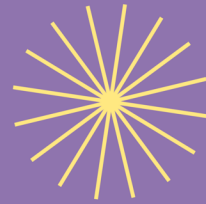


Increase involvement of underrepresented individuals in AI research

06



Collaboratively develop a "Debiaser technology" to address biases in decision making



What is the Debiaser



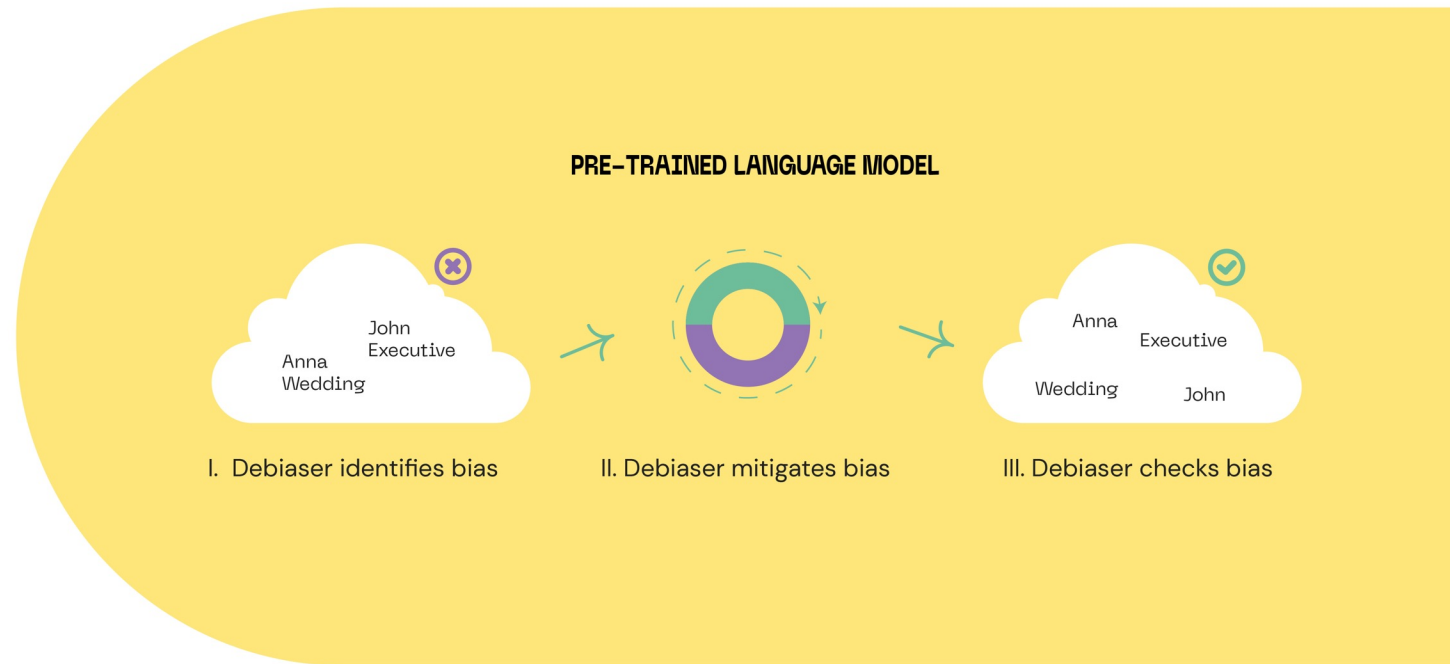
Funded by
the European Union

BIAS



DEBIASER

- The latest NLP technology is based on so-called pre-trained language models. Being trained on a large corpora of text, such models provide numerical representations of words, making it easier for the computer to handle human language.
- **The Debiaser** will provide a toolkit for identifying and mitigating biases in such language models, making them safer to apply in Human Resources Management.





During the recruitment phase, the Debiaser shows which issues of the applications might be subject to discrimination and how to deal with them for bias mitigation



Provides human-understandable explanations on the parts that impacted the automated decisions proposed to recruiters, building trust in them.

The challenge of bias in AI-driven recruitment starts with the definition of what fairness means. **The Debiaser creates a use-case specific definition of fairness, ensuring similar candidates are treated in a similar way.**

Based on the BIAS research, the Debiaser will support identifying and building the domain – and organisation – specific knowledge necessary to implement the Case-based Reasoning* and conducting regular checks to ensure fairness and consistency in the recruitment process.



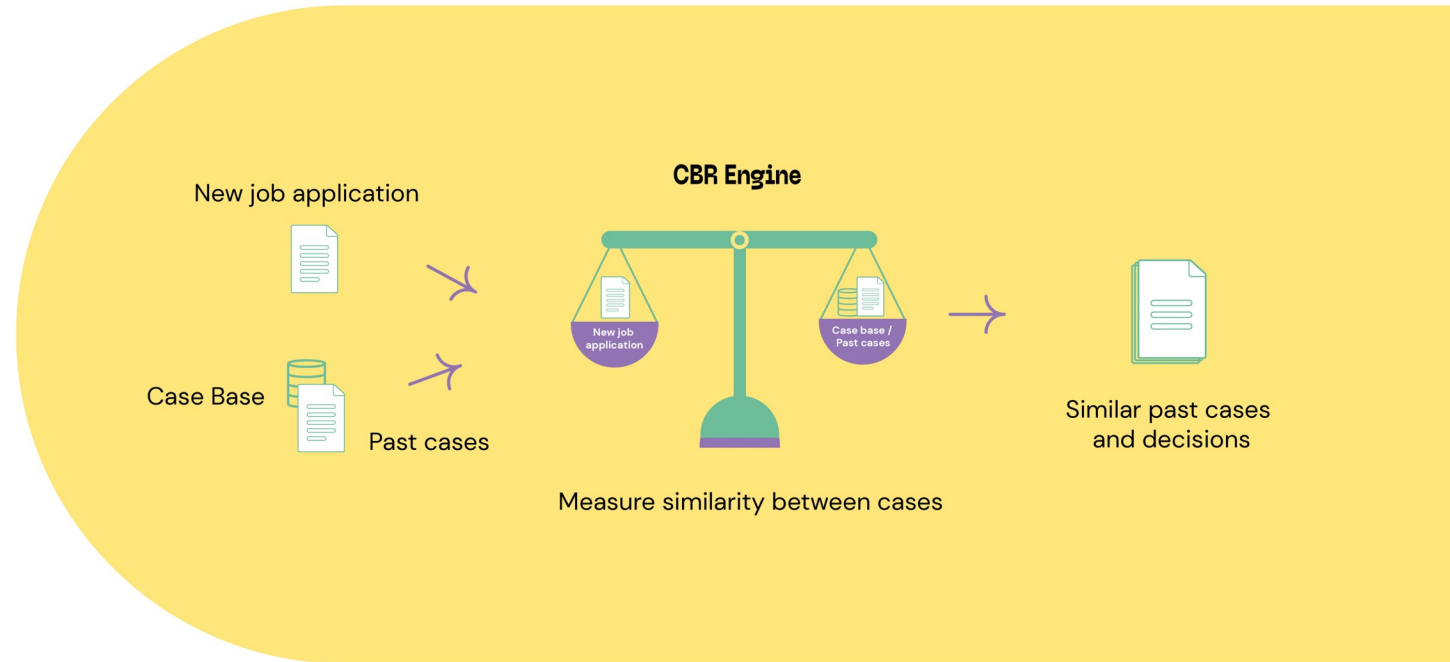
DEBIASER

*CBR solves new problems by reusing successful solutions previously applied to similar problems.

It comprises 2 components:

- A **Case Base** that contains the problems solved in the past.
- A **CBR engine** that retrieves similar past cases to the new one and reuses their solutions for solving the new problem.

This can ensure "individual fairness" since similar individuals will be evaluated in a similar way.





Target audiences of the BIAS project



Funded by
the European Union

BIAS



AUDIENCES

Key players



- Platforms using AI systems to analyse data
- Technology developers
- Public and private investors

Context Setters

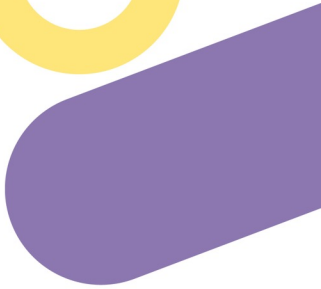


- Local, national and international policy makers
- Standardisation organisations
- Professional networks and recruitment platforms

Advocates



- Labour unions and individual workers
- Educators
- Citizen groups and advocacy organisations that promote gender equality and diversity
- Academics, researchers, and think tanks





They can benefit from the project by participating and being part of:

- National Labs
- Interviews and surveys
- Capacity building and awareness raising sessions
- Co-creation workshops
- Policymaking activities
- Trustworthy AI helix
- The Debiaser open-source software



What the BIAS project has to offer



Funded by
the European Union

BIAS



RESEARCH AND IMPACT



The **research methodology** is based on the following pillars

01

The creation of national labs in each country (communities of practitioners, employees, HRM, and AI specialists with a special focus on underrepresented communities). Members of the National Labs will participate in needs analysis and stakeholder involvement through surveys, interviews and co-creation workshops



RESEARCH AND IMPACT



02

AI research and development with a focus on Natural Language Processing (NLP) and Case-based Reasoning (CBR)

03

The creation of the Debiaser, our proof-of-concept technology with modules that both identify and mitigate bias and unfairness in decision making, that will be made available to the AI community



RESEARCH AND IMPACT



04

Ethnographic fieldwork with employers, employees, and AI developers from different European countries providing information about current experiences and future scenarios of the BIAS model and AI in employment settings to PhDs and researchers



RESEARCH AND IMPACT

4 months of fieldwork in each country, in companies/organisations utilising or developing AI for assessment for hiring and work.

COUNTRY	DEVELOPERS	EMPLOYERS	EMPLOYEES
● ICELAND		10	25
● NORWAY		20	50
● ITALY	50	20	50
● NETHERLANDS		20	50
● TURKEY		20	50
TOTAL (365)	50 PAN EUROPEAN	90	225





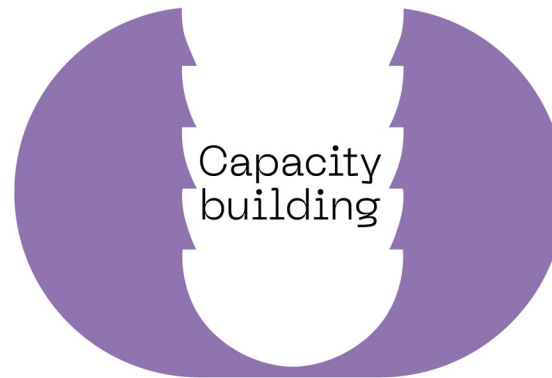
ACTION PLAN



These strands will be combined with a detailed and extensive impact strategy consisting of:



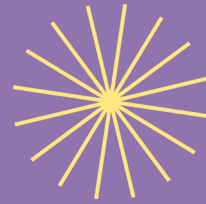
On the importance of tackling gender and intersectional biases in AI



To equip the AI and HRM community with tools to prevent bias in AI



That companies can use to reduce biases in their HR practices



Who we are



Funded by
the European Union

BIAS



The BIAS project is composed of 9 partners (from 9 different countries) from:

- AI/NLP solutions and research;
- SSH (Social Sciences and Humanities) knowledge;
- Diversity and inclusion in organisations in HR practices;
- Dissemination and communication;
- Industrial uptake and commercialization.

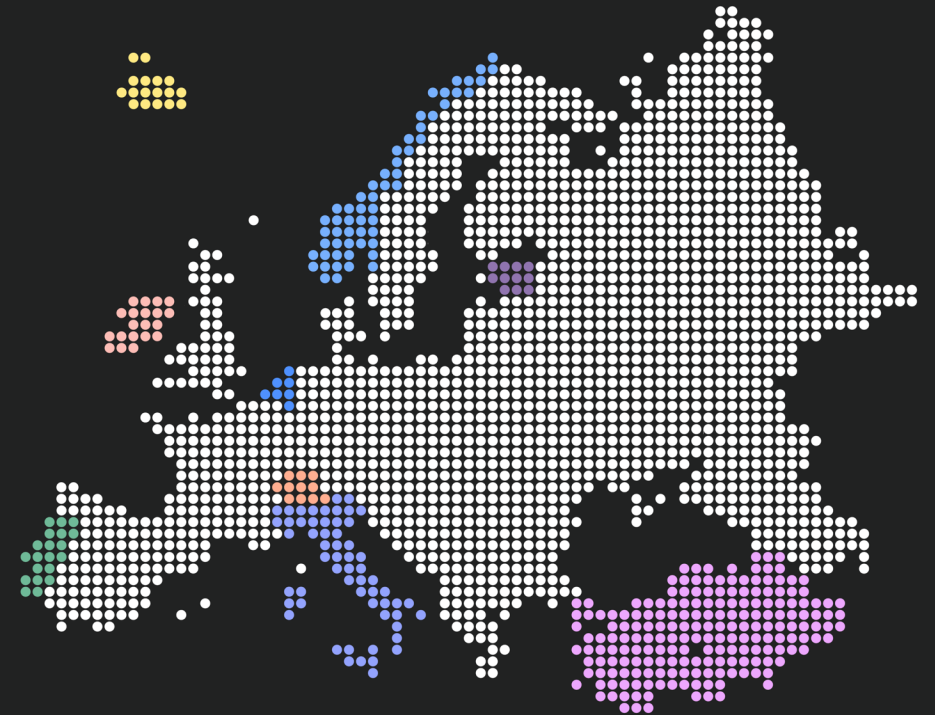


Our consortium is made up of **4 universities, 3 communication partners, 1 large industrial organisation** and **1 small and medium enterprise**:

BIAS

CONSORTIUM MAP

- NORWAY – NTNU
- SWITZERLAND – BFH
- ICELAND – HI
- PORTUGAL – LOBA
- IRELAND – CHX
- ITALY – SVEN
- NETHERLANDS – ULEID
- ESTONIA – DIGI
- TURKEY – FARPLAS



Funded by
the European Union

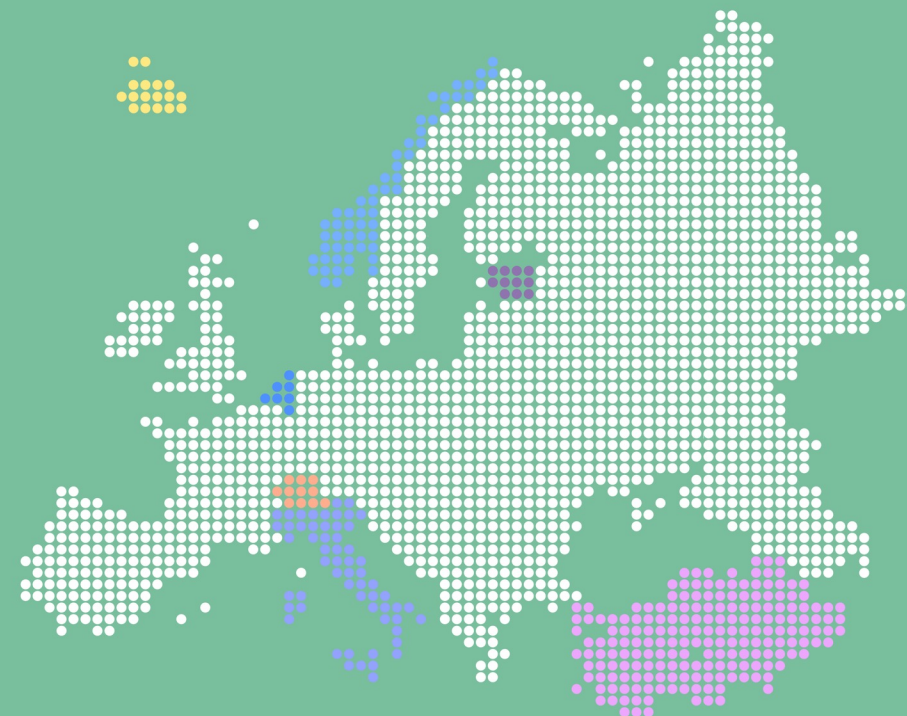
BIAS

Would you like to shape a future where fairness and inclusion thrive? Take the leap and apply now at biasproject.eu/nationallabs/

BIAS
NATIONAL LABS

- **NORWAY** – NTNU
- **SWITZERLAND** – BFH
- **ICELAND** – HI
- **ITALY** – SVEN
- **NETHERLANDS** – ULEID
- **ESTONIA** – DIGI
- **TURKEY** – FARPLAS

You have **National Labs** awaiting you in the following countries:



BIAS

Mitigating biases
of AI in the
labour market

Consortium



NTNU



UNIVERSITY
OF ICELAND

LOBA®



CrowdHelix



SMARTVENICE_{SL}



Universiteit
Leiden
eLaw



DigiTouch

farplas



Bern University
of Applied Sciences



Funded by
the European Union

@BIASProjectEU

