"Designing a responsible AI assisted hiring pipeline through alignment between recruitment, software development and interdisciplinary approaches",

Reflections from FINDHR panel debate, 16 October 2025 by Francesca Maddii

The FINDHR project (Fairness and Intersectional Non-Discrimination in Human Recommendation) was created to address a simple yet critical question: how do we ensure that Artificial Intelligence hiring technologies serve all applicants equally? At the closing conference computer science and recruitment experts reflected on the lessons learned from this three-year interdisciplinary project in the panel debate titled: **"Combining Innovation in Software Design and Human Resource Management to Reduce Discrimination in the AI-Assisted Hiring Pipeline".**

The panellists explored how fairness to reduce discrimination can and should be built into every stage of AI-assisted hiring, tackling the technical, social and legal strategies to build a fairer and responsible recruitment process. They agreed that AI hiring tools can reduce discrimination only when humans remain accountable for their design and implementation, with a commitment to tackle biases and discriminatory processes where they arise.

Panel Discussion: "Combining innovation in software design and human resource management to reduce discrimination in the AI assisted hiring pipeline"

From right to left: Anna Gatzioura, Asia Biega, Anna Via and moderator Frederik Zuiderveen Borgesius

Anna Gatzioura, Postdoctoral Researcher, Universitat Pompeu Fabra and FINDHR researcher.

Asia Biega, tenure-track faculty member (W2) leading the Responsible Computing group at the Max Planck Institute for Security and Privacy (MPI-SP). She is also FINDHR principal investigator.

Anna Via, Head of Artificial Intelligence at InfoJobs, the leading job board in Spain and industry partner to FINDHR.

Moderator: Frederik Zuiderveen Borgesius, Professor of ICT and law at Radboud University, At Radboud University he is affiliated with the iHub, Radboud University's interdisciplinary research hub on Digitalization and Society. He is a FINDHR researcher.

Synthetic data and fairness

Anna Gatzioura kicked off the panel debate by presenting the UPF team's contribution to the FINDHR project. A synthetic dataset of CVs has been created to be a fairness benchmark in AI-based hiring systems. Synthetic data is suggested by the EU AI Act as a methodology for training AI given that sensitive data of existing people is protected by the General Data Protection Regulation (GDPR) and therefore very time-consuming to access in a legally compliant way for testing.

The team at UPF collected and anonymized donated European CVs, then generated synthetic versions using a mix of manual and automated techniques to ensure realism and diversity, producing about 2,000 CVs covering a variety of sectors.

The research is expected to contribute to standardization efforts, and she emphasized that future AI standards must center on protecting intersecting rights of individuals and apply across all stages of the AI lifecycle. A key takeaway from the research is that ensuring fairness in AI-based hiring is a combination of stress testing a system before launch, but also proper monitoring once it is in production. "*We recommend the use of synthetic data for benchmarking in the pre-deployment phase, as once something is deployed in production, it's very difficult to predict the evolution of the system and to correct what will happen*", she noted.

Monitoring an AI system when being used draws attention to the important role of real data in this process. The data needs to reflect the current society, showcasing what the workforce looks like to achieve the best results. A tool will interact differently when it is being used in the market of a different segment and region to where it was tested for.

Overall, Gatzioura explained that synthetic data needs to be of a good quality in order to be useful: if the original dataset is not representative, the synthetic data will not be either.

Mapping and Correcting Bias

Biases caused by the kind of data being processed is one type of bias that exists. Asia Biega, of the Max Planck Institute for Security and Privacy, researched with her team the occurrence of bias in algorithms from a software design point of view, observing the causes and what can be done to mitigate it. She described three layers of bias, which can be divided into technical bias (resulting from model design or optimization choices), pre-existing bias (derived from societal inequalities reflected in the data), and emergent bias (when systems interact with new and unpredictable contexts).

Her work in FINDHR focused on developing a repository that includes tools to track algorithmic fairness over time, including a multiparty computation protocol (a guideline for collecting and storing sensitive data in a distributed way), and explaining how decisions are made.

Biega stressed that the point made by Gatzioura that bias can emerge at all times during the system's lifecycle as a fundamental learning. In this context she found through the research that the "devil is in the details": a solution can seem clear in theory, but the way it gets technically implemented can be (too) difficult. *"We can't assume we solved a problem after proposing a solution; there are so many steps in between".* With this she reflected on the multiparty computation protocol (MPC) that was developed as part of FINDHR. MPC is a computation scheme adopted by two or more parties (e.g. two organisations) to compute a function of their data without revealing that data to other parties. It is on paper a very clear framework but implementing it brings with it, its own challenges that can lead to other preferred solutions.

Finally, a technical solution should always account for the human factor: people should be part of the conversation and be aware of the processes that are happening within the box of the algorithm to make informed decisions and avoid the creation of more bias.

Covering the gap between research and real-world application

Anna Via, Head of Artificial Intelligence at InfoJobs a leading recruitment company in Spain, brought an industry viewpoint to the discussion. Her company worked with FINDHR to implement the toolkit in a real recruitment environment.

InfoJobs developed two demonstrators: the first is *a discrimination monitoring dashboard,* visualizing indicators such as exposure, diversity, and bias across different job categories. The second dealt with fair ranking interventions, focused on modifying the algorithm that orders candidates to evaluate how fairness adjustments affect the overall ranking. The results of using these demonstrators were highly encouraging, and the internal reaction within InfoJobs was positive: sales, product teams, and clients all expressed strong interest in being able to quantify fairness and diverse outcomes. Larger client companies, which often have dedicated diversity officers, also showed enthusiasm for tools that would make ethical hiring practices more measurable and clearer.

Given these outcomes, Anna Via expressed that research that FINDHR has promoted is deeply relevant for today's hiring processes, as companies need and want these AI tools to operate better and avoid further discrimination. Because, as both Anna Via and Anna Gatzioura mentioned, AI tools that are biased and not trustworthy thus leading to discrimination in the hiring pipeline present: "*a lose-lose scenario: companies miss out on capable and skilled employees, and candidates cannot find a job that would match their qualifications*".

Via also warned against "implementation paralysis", waiting for perfect systems before taking any action, as this is not an achievable goal. She explained that observing how others use AI can help shape good practices and avoid mistakes, with as objective, making sure that the algorithms implemented are being used responsibly.

Anna Via finally reflected on the tension between the wish of a company or recruiters to have a practical set of rules that it can follow to manage its AI and the reality that AI tools are too complex and unknown (at the moment) to have such rulebooks in place for each context and AI system. However, companies can improve the management of their AI, and starting with auditing can often be a good entrypoint.

Need to bridge disciplines

All speakers agreed that interdisciplinary and multistakeholder collaboration is difficult, but indispensable.

Asia Biega described aligning stakeholders across sectors as one of the keys to success. This starts with the design; when teams map out what each stakeholder values, future cooperation becomes easier. Anna Via added that this requires interdisciplinary communication. Between HR, law, and tech, there are many different actors at play, but to implement good practices, there needs to be an active conversation between them all, as they all influence and can help each other.

Anna Gatzioura also pointed to organizational and political pressures beyond the control of teams, as sometimes developers get blamed for bias, but deadlines and priorities are set much higher up. Frederik Zuiderveen Borgesius pointed out that part of interdisciplinary communication should be a deep reflection on the words we use. We all often use the same words, such as fairness, discrimination, and transparency, but we mean different things, and the only way to understand each other is to invest in communicating with each other.

Audience interventions

The Q&A section raised challenging questions about who is accountable for biased outcomes and how fairness can be measured.

The speakers in response agreed with each other that humans remain legally and morally responsible for AI outcomes, and that monitoring and intervention, while separate modules in a system, feed into each other in a feedback loop. In terms of which bias metrics to use, the answer, based on the FINDHR research context, is to decide this depending on the context. FINDHR started mapping and developing all kinds of applicable metrics. A catalogue of metrics should be further developed to understand bias and measure discrimination, which would need the collaboration of all actors: domain experts, industry practitioners, and legal scholars. Companies should be incentivized to implement practices to mitigate discrimination, as there are many benefits: good PR, long-term trust building with users, and sustainability.

The context influences the technologies that are most practical to use, for example for deciding to use the multiparty computation protocol. In some countries, it is common to provide certain personal characteristics, while in others this is not common. These cultural expectations are not easy to change; at least this requires not only a technical solution.

Also, the application frameworks are different depending on the job that people are applying to, noting especially differences between blue-collar and white-collar ones. As such, the bias that is encountered will also be different. This sparked the question: should we fundamentally change the way we apply to jobs, with what we know now about AI and algorithm-based hiring?

Key Takeaways and Recommendations

Throughout the duration of the panel, many relevant suggestions and insights were shared. These were the main takeaways:

- Reducing discrimination in AI assisted hiring processes is a shared responsibility across developers, HR, and policymakers. Collaboration across disciplines leads to better outcomes.

- Fairness must be built into the algorithm from the start (preferably within an interdisciplinary team), and developers must consider societal values and context.

- Synthetic data can be a safe and effective tool for stress-testing on discrimination when developing an AI, given that the data is drawn from good quality data.

- Companies should view reducing discrimination in AI-assisted decision-making not solely as a compliance issue, but as a trust-building mechanism to draw in and keep customers; transparency and explainability are essential to build said trust.

- Continuous monitoring is needed to keep systems fair when being used in the market.

Closing Reflections

In his closing remarks, Frederik Zuiderveen Borgesius reflected on the tension between data privacy and fairness. While there are good reasons not to collect certain data, we can't always tell whether discrimination is happening without it. He also urged researchers to communicate more openly with policymakers and the public.

The session ended with optimism, as AI has proven to be a useful tool when used responsibly from the perspective of Infojobs. It requires the joint work of developers, HR professionals, and policymakers among others to incorporate AI tools in the hiring pipeline that works towards equal opportunities and reduces anti-discrimination practices.

In a fast-paced and ever-changing job market, the task is to keep raising awareness and understanding on the complexities surrounding the development and use of AI among all the stakeholders involved in the recruitment process. As Anna Via concluded, we do not need to stop using AI in recruitment, but make sure it gets used responsibly.

Written by: Francesca Maddii. Francesca is a Young Professional at WIDE+, following her graduation as a master's student in Communication Strategies at the University of Padova, where she focused on Political Communication and Gender Disparities. At WIDE+, she is involved in the communication efforts of WIDE+ and assists in multiple active projects.

About FINDHR. It combines legal, ethical, and technical resources for all groups involved in AI assisted hiring processes, in particular employers, policymakers, and software developers. It

contributes to making fairness to some extent measurable and actionable, turning academic research into tools that can help to reduce discrimination in hiring when using algorithms: see: www.findhr.eu.

"Designing a responsible AI assisted hiring pipeline through alignment between recruitment, software development and interdisciplinary approaches", Reflections from FINDHR panel debate,16 October 2025

The FINDHR project (Fairness and Intersectional Non-Discrimination in Human Recommendation) was created to address a simple yet critical question: how do we ensure that Artificial Intelligence hiring technologies serve all applicants equally? At the closing conference computer science and recruitment experts reflected on the lessons learned from this three-year interdisciplinary project in the panel debate titled: **"Combining Innovation in Software Design and Human Resource Management to Reduce Discrimination in the AI-Assisted Hiring Pipeline".**

The panel was moderated by Professor Frederik Zuiderveen Borgesius of Radboud University. The panel featured Dr. Anna Gatzioura (Universitat Pompeu Fabra), Dr. Asia Biega (Max Planck Institute for Security and Privacy), and Anna Via (Head of AI at InfoJobs). They explored how fairness to reduce discrimination can and should be built into every stage of AI-assisted hiring, tackling the technical, social and legal strategies to build a fairer and responsible recruitment process. The panellists agreed that AI hiring tools can reduce discrimination only when humans remain accountable for their design and implementation, with a commitment to tackle biases and discriminatory processes where they arise.

Key Takeaways and Recommendations

Throughout the duration of the panel, many relevant suggestions and insights were shared. These were the main takeaways:

- Reducing discrimination in AI assisted hiring processes is a shared responsibility across developers, HR, and policymakers. Collaboration across disciplines leads to better outcomes.
- Fairness must be built into the algorithm from the start (preferably within an interdisciplinary team), and developers must consider societal values and context.
- Synthetic data can be a safe and effective tool for stress-testing on discrimination when developing an AI, given that the data is drawn from good quality data.
- Companies should view reducing discrimination in AI-assisted decision-making not solely as compliance issue, but as a trust-building mechanism to draw in and keep costumers; transparency and explainability are essential to build said trust.
- Continuous monitoring is needed to keep systems fair when being used in the market.

Anna Gatzioura stressed that a key takeaway from the research is A key takeaway from the research is that ensuring fairness in AI-based hiring is a combination of stress testing a system before launch, but also proper monitoring once it is in production. "*We recommend the use of synthetic data for benchmarking in the pre-deployment phase, as once something is deployed in production, it's very difficult to predict the evolution of the*

system and to correct what will happen", she noted. Gatzioura explained that synthetic data needs to be of a good quality to be useful: if the original dataset is not representative, the synthetic data will not be either.

Asia Biega stressed that the point made by Gatzioura that bias can emerge at all times during the system's lifecycle as a fundamental learning. In this context she found through the research that the "devil is in the details": a solution can seem clear in theory, but the way it gets technically implemented can be (too) difficult. "*We can't assume we solved a problem after proposing a solution; there are so many steps in between*".

Anna Via, Head of Artificial Intelligence at InfoJobs a leading recruitment company in Spain, brought an industry viewpoint to the discussion. Her company worked with FINDHR to implement the toolkit in a real recruitment environment.

Anna Via expressed that research that FINDHR has promoted is deeply relevant for today's hiring processes, as companies need and want these AI tools to operate better and avoid further discrimination. Via also warned against "implementation paralysis", waiting for perfect systems before taking any action, as this is not an achievable goal. She explained that observing how others use AI can help shape good practices and avoid mistakes, with as objective making sure that the algorithms implemented are being used responsibly.