This is a draft chapter. The final version will be available in *How to Manage International Multidisciplinary Projects*, edited by Linda Hantrais, forthcoming 2022, Edward Elgar Publishing Ltd.

# 5. Responsible AI for labour market equality (BIAS)

**Alla Konnikov, Irina Rets, Karen D. Hughes, Jabir Alshehabi Al-Ani, Nicole Denier, Lei Ding, Shenggang Hu, Yang Hu, Bei Jiang, Linglong Kong, Monideepa Tarafdar and Dengdeng Yu**

PROJECT DESCRIPTORS

**Duration**: 1 February 2020−30 July 2023 (3.5 years)

**Web Links**: https://www.lancaster.ac.uk/lums/research/areas-of-expertise/centre-for-technological-futures/responsible-ai-for-labour-market-equality/

**Principal Investigators**: Linglong Kong, University of Alberta, Canada (mathematical & statistical sciences); Monideepa Tarafdar, University of Massachusetts Amherst.

**Co-Investigators**: University of Alberta, Canada − Nicole Denier (sociology); Karen D. Hughes (sociology & business); Bei Jiang (mathematical & statistical sciences)

University of Essex, UK − Hongsheng Dai (mathematical sciences); Berthold Lausen (mathematical sciences)

Lancaster University, UK - Yang Hu (sociology); Bran Knowles (computing & communications)

**Postdoctoral Fellows and Graduate Students**: University of Alberta, Canada − Lei Ding (mathematical & statistical sciences); Wenxing Guo (mathematical & statistical sciences); Alla Konnikov (sociology); Meichen Liu (mathematical & statistical sciences); Jinhan Xie (mathematical & statistical sciences)

University of Essex, UK − Jabir Alshehabi Al-Ani (mathematical sciences); Shenggang Hu, (mathematical sciences)

Lancaster University, UK − Irina Rets (management science)

University of Texas, US − Dengdeng Yu (mathematics)

**Funding Organisations**: Economic and Social Research Council (ESRC) and Social Sciences and Humanities Research Council of Canada (SSHRC), funding programme Canada−UK Artificial Intelligence Initiative: Building competitive and resilient economies through responsible AI; award number ES/T012382/1, amount £508,000.

## THE BIAS CASE STUDY

The BIAS case study describes the aims and objectives of the 'BIAS: Responsible AI for labour market equality' project, the research process, and the lessons learned at the mid-way point of the project. It draws on two main sources of information: interviews conducted internally with six founding members of the project in August 2021, and a survey of the full team that was conducted in September 2021 to gather information on the backgrounds, perspectives, and experiences of all team members who were currently active in the project. Konnikov, Rets and Hughes led the study

conceptualisation, instrument design, data analysis, and writing-up activities. Konnikov and Rets led on the data collection and administration of the internal interviews and the survey with the other project team members. The remaining authors (in alphabetical order) contributed to the case study as it developed; they read, edited and approved the final text of the case study.

## Aims and Objectives

An overarching aim of the BIAS project was to ensure a genuinely interdisciplinary approach. BIAS had two primary aims:

- To refine existing understandings of AI-driven bias, using an interdisciplinary, cross-national, multi-stakeholder approach
- To develop 'responsible AI' that can potentially mitigate gender and ethnic biases in algorithmic labour markets, using Canada and the UK as test sites

## Themes and Research Questions

BIAS is an international interdisciplinary project that seeks to understand the role of artificial intelligence (AI) in reproducing gender and ethnic biases in labour market processes, such as job postings and hiring, that are increasingly digitalised. The project was motivated by the 'digital turn' in contemporary labour markets and growing concerns over the role of AI in reproducing and exacerbating market inequalities. The BIAS project speaks directly to several themes and national priority agendas in Canada and the UK, including the rise of AI and gender and racial disparities in hiring, pay gaps and innovative practices, since both countries embrace digital transformations as part of their economic and industrial strategies.

The project was designed to generate theoretical, empirical, applied, and policy-related contributions by examining four research questions:

1. What are the sources and dimensions of gender and ethnic bias in labour market processes?
2. How do different parties in the labour market perceive and respond to bias, and how do biases lead to labour market inequalities?
3. How are biases currently reproduced in AI algorithms and through AI human interactions?
4. What approaches and techniques can be used to minimise and mitigate such biases?

# THE RESEARCH PROCESS

The project's key objectives, research design, and team composition were shaped by the funders' requirements in the formal call for proposals, and by the expertise and interests of the founding team members. This case study examines the decisions that were taken and the range of factors that came

into play as the project evolved, focussing on the international and disciplinary dimensions of the study.

## Meeting Funders' Requirements

The call for proposals from the Canada-UK Artificial Intelligence Initiative played an important role in shaping the aims and objectives of the project. A central goal of the initiative was to spark more interdisciplinary research on AI to aid the development of more inclusive and responsible AI technologies. Three requirements in the call determined the project design and objectives from the outset, including the involvement of at least two distinct research domains: the participation of researchers from Canada and the UK, and an explicit focus on the development of 'responsible AI'.

Equally important for shaping the project were the expertise and interests of the founding project team. While topics such as smart cities, healthcare, and democratic governance were highlighted in the call for proposals, the issue of labour markets received only a brief mention. Yet, the social science team members had strong interests in labour markets and human resource management, and the growing role of AI in reproducing gender and ethnic bias. The team members in the mathematical and computational sciences were seeking new 'areas of application' for their research, and the issue of gender and racial bias seemed interesting and challenging.

The grant proposal was well received, and was one of the ten projects funded under the initiative (8.8% success rate). Funding came from the two national councils, in roughly equivalent amounts, with sufficient funds to cover project costs, including post-doctoral and student funding, data and software. Institutions in both countries also provided basic support through technology infrastructure, office space, and some administrative assistance.

## Building the Project Team

While the funding call shaped the team composition, with respect to disciplinary and national location, equally important were pre-existing professional relationships developed through conferences and events, and from working at the same institution. Some connections went back many years, to graduate school colleagues. Other were more recent: for example, colleagues meeting at an event organised by their university to spark interdisciplinary research.

In the initial stage of the project, the team was assembled at two UK universities where faculty members had prior connections (Lancaster and Essex). While they shared their interests in AI, they were working in distinct, albeit inter-related, disciplines, including management science and information systems, sociology, computing and communications, mathematical sciences and computational statistics. One of the UK team members then contacted a colleague in Canada (Alberta) to gauge interest in a collaboration. A Canadian team was assembled quickly across three disciplines: mathematical sciences and computational statistics, sociology and organisational studies.

While some prior connections existed within each disciplinary field, the recruitment process required actively seeking out individuals with the requisite mix of disciplinary knowledge, experience, and interests.

The final founding team included nine researchers representing two countries, three universities, and five different disciplines. Together they had educational and work experiences spanning six countries: Canada, China, Germany, India, the UK and the United States. Once funding was awarded, the project team expanded further with the hiring of postdoctoral research fellows and graduate students as research assistants.

Project leadership was provided by two principal investigators, one in each country: Monideepa Tarafdar in the UK and Linglong Kong in Canada. The team did not employ a professional project manager, but a small amount of administrative support was provided through Lancaster University. Much of the academic and collaborative work was accomplished through a 'self-managing team' approach with several team members fulfilling an important 'bridging' role. The project consisted of 18 researchers at its inception, though several graduate researchers ceased to be involved when they completed their formal programmes. It also included an advisory board of ten experts from industry and academia who provided advice and support as needed, and attended an annual meeting with all team members.

Beyond the requisite mix of disciplinary skills and backgrounds, one additional characteristic that proved to be important was an 'interdisciplinary mindset', together with interdisciplinary experience from past projects, or an active interest in related disciplines. Confirming the importance of this characteristic, the internal survey of team members found that 30.7% had 'significant' interdisciplinary experiences, and 46.1% had 'some' experience; 23.1% had no interdisciplinary experience. Team members with an interdisciplinary background had an increased appreciation of, and interest in, the overarching objectives of the project, and understood how the different approaches, concepts and methods involved might produce novel insights. They also understood the challenges that might arise when working across disciplinary boundaries, and how to resolve them productively (Fiore et al., 2008).

## Research Design and Methods

The BIAS project utilises a new model of international interdisciplinary AI research that involves mixed/multi-method, co-produced AI research. The project was deliberately comparative and interdisciplinary. It used two interconnected work packages (WPs) to bring together different disciplines to pool national data and analytic techniques, using labour force and job platform data; to undertake qualitative interviews with platform participants at the organisational and individual level; to model, simulate and refine understandings of bias; and to develop and test AI designed to mitigate the biases detected. WP1 involved an algorithmic-oriented examination of bias, including data mining

to detect bias, and to test new AI algorithms for bias mitigation. WP2 adopted a multi-stakeholder perspective (employer, employee, industry), and qualitative investigations, to understand AI use in processes of job advertising, screening and hiring.

Qualitative research studies with recruitment companies and employers were undertaken firstly in the UK and then in Canada in 2022, using a complementary research design. Data collection and engagement with job seekers were planned for the final year of the project as well as knowledge exchange activities with policy researchers, developers and computer scientists, and organisations involved in regulating the use of AI.

From the outset, work tasks were divided across smaller intra-disciplinary units, enabling collaboration to take place both within and between these smaller units, and then with the larger team. Results from the internal survey confirmed the effectiveness of this strategy, with the highest level of satisfaction for team members being associated with intra-disciplinary communication within the small teams. One of the investigators described working in these teams as particularly important for maintaining conceptual and methodological rigour in all disciplines: '…we meet in the bigger group for the bigger picture, but we also need to meet in the smaller groups, because each discipline has to have their own excellence there as well.'

The aim of ensuring that the project would have a genuinely interdisciplinary approach had a number of implications for the research processes and for the professional development of the scholars involved. A commonly held view in the project team was that the absence of any one of the project's constituent disciplines would constrain the whole process. Although in part a requirement of the grant, the project's interdisciplinary design enabled a more comprehensive and refined account of a complex problem such as bias, which can be understood at different levels (individual, cultural, personal, organisational, or structural), and as action or outcome. Having diverse disciplines involved ensured the necessary theorisation and contextualisation for the research problem, supported by the knowledge exchange and interdependence between the intra-disciplinary teams.

Reaching agreement in interdisciplinary work can often pose challenges, a point that was highlighted in the interviews with team members. A particular example concerned the definition and operationalisation of the central concept of the project: 'bias'. In early meetings, it became evident that different disciplines defined, understood, and operationalised bias in different ways. One investigator illustrated the challenge involved: '… at the beginning, we actually struggled a lot. We asked the sociology and the management team to define bias. And for our AI team, we asked the same question. And we couldn't match each other, even in the same team, each definition was different.' To reach agreement on the conceptual definition of bias, the team conducted an exercise where each investigator wrote down their own definition of bias and then shared it with the rest of

the team. This exercise enabled the whole team to gain a better understanding of different definitions and provided a starting point for future discussions.

In team meetings, the aim was not to achieve 'sameness' by coming up with one overarching understanding of bias, but rather to gain a better understanding of the ways in which bias was conceptualised and studied in each discipline (Fiore et al., 2008). This approach highlighted the dynamic nature of interdisciplinary knowledge exchange. The initial understanding of the main concepts in the research underwent transformations resulting from the contributions of each discipline. Remaining open to the incorporation of different views about the conceptual understanding of the phenomena of interest helped to bridge, rather than resolve, the conceptual interdisciplinary differences.

The meetings offered a platform for dialogue; they helped team members get to know one another, and to learn about the concepts, meanings and methodologies in different disciplines. To facilitate the process of reaching agreement on conceptual and methodological issues, team members developed the following strategies: learning about conceptual differences; embracing conceptual differences; and enabling conceptual and methodological understandings to evolve and change over time. The small intra-disciplinary meetings facilitated in-depth, rigorous discussion of concepts , while the small interdisciplinary meetings created conceptual bridges between the teams.

The interdisciplinarity of the research design also had implications for the innovative use of research methodologies. Since each project discipline was grounded in particular epistemologies, the mix of disciplines allowed for the integration of the methods associated with each discipline by drawing on their individual strengths (Fiore et al., 2008). Several team members felt that their research inquiries were expanded by the multidisciplinary methodologies embraced in the project. As one team member noted, instead of 'waiting for the next five-ten years' to import the methods from other fields into their analysis, the interdisciplinary collaboration meant that, together, the team was 'innovating for the next round of methods, in a fast-breaking way'.

The need to prevent potential misunderstandings increased interdependence between the different intra-disciplinary teams. They used visual tools, such as pictures, graphs and figures, offering a universal language that scholars from the different disciplines could all understand, to summarise ideas and concepts, and translate disciplinary terminology in a more 'research universal' and accessible way. Interdependence was developed between the different intra-disciplinary teams by creating shared tasks and fostering adaptivity to change, as the project's conceptualisations and methodologies evolved over time.

Some team members explained how what they initially envisioned as the primarily independent work of intra-disciplinary teams, with occasional collaboration, turned out differently as the project

progressed. The complexity involved in conceptualising and operationalising bias sparked more frequent exchange. For example, the computational team found that, by collaborating closely with the sociology team, they were able to solidify concepts and methods, and make their analysis more meaningful. This process informed work on a paper where the sociology team took the lead in developing a list of gender-related words in job advertisements, which the computing team then used to refine their approach to gender bias mitigation (Hu et al., 2022). As explained by a computing team member: 'Although we could, probably, rely on common sense, we don't have to in this project. We can actually use input from the experts from other disciplines that can help us do a better job with finding the gender words, which will allow us to define gender bias more precisely.'

The team members highlighted two major approaches to interdisciplinary collaboration that were embraced as a part of the team culture in designing the research: openness towards various intellectual and methodological traditions; and 'flat', egalitarian, divisions of labour within the team. Several strategies to support cross-disciplinary communication also emerged.

Team members described openness as an approach that facilitated greater flexibility, acceptance and dialogue between participants trained in different disciplines, explained by one team member as a constant 'willingness to be out of one's comfort zone'. Pairing an openness to diverse intellectual interests and goals, with a degree of compromise, was an essential strategy in a complex, interdisciplinary and international project. As observed by another team member: 'We all had to carve out and include and exclude bits of our interests to make this collaboration work.'

This culture of flexibility and openness shaped the division of labour within the team revealing a particularly egalitarian approach. Despite including academics of various ranks, and with diverse specialties, the team's organisational and communication structures remained flat. Close attention was paid to members' unique interests and capabilities in assigning tasks and responsibilities. The internal survey with team members highlighted how a flexible division of labour had helped them to engage in activities that they felt most passionate about and most capable of pursuing. Team members also stressed the evolving nature of their roles over time, as some initial roles were revised and adjusted. One example of this adjustment was the evolving 'bridging' role played by team members who attended meetings of multiple intra-disciplinary teams and helped to facilitate the connections between them.

Over time, team members developed a combination of formal and informal communication strategies to promote understanding of specific issues and facilitate the process of reaching agreement. Ongoing internal communication practices included monthly interdisciplinary meetings with all team members present, and weekly and bi-weekly intra- and inter-disciplinary meetings of smaller teams working on similar tasks (Tarafdar & Davison, 2018). Larger team meetings offered a valuable platform for resolving ambiguities and misunderstandings, and for reaching consensus on work

packages and processes. Virtual meeting technologies were especially helpful not only during the pandemic, but also for building connections within the team in the early stages, given the physical distance between the institutions involved in the project.

Within a large, interdisciplinary and international team, where collaboration unfolds remotely and involves a significant portion of email correspondence, the potential for misunderstanding is high, and the risks of accumulating problems are considerable. The team members consistently emphasised how the meetings could be used to resolve issues in a timely manner and improve the understanding of each other's perspectives. Smaller group meetings, typically intra-disciplinary, were rated in the survey as particularly effective for knowledge exchange and for clarifying the concepts and operationalisations practised in each discipline. Team members ranked the small group meetings as a top communication strategy and referred to them as the main facilitators of the working process. Flexible on-demand meetings were the second most highly-rated communication strategy. In the words of one team member, they 'expedited the work rhythm and cut the long chain of formalities'.

In addition to the positive roles of openness and the intra- and interdisciplinary meetings, team members also used other strategies to minimise misunderstanding and the mistranslation of concepts across disciplines. Each discipline used a different 'language' (terminology) that had to be 'translated' across the teams (Monteiro & Keating, 2009). According to one team member: 'We had to explain things, not using our own language. We had to explain things to make sure that a layperson can understand it.' Some team members took on responsibility for bridging the communication gap between the small inter- and intra-disciplinary group meetings by actively translating the concepts. These roles were naturally taken by the members of the team who 'fall in between the gaps that cover slightly different areas' and have a solid understanding of the concepts and methods used in different disciplines.

With respect to the professional development of the scholars involved in the project, interdisciplinarity encouraged each researcher to perceive the familiar research methods that they commonly used in a new light. In the internal survey, team members ranked the development of new skills and new ways of intellectual thinking through interdisciplinary work as highly impactful for their career advancement. As one team member explained: 'I'm a quantitative sociologist. It's [the approach used in another discipline] not something entirely new to me, but I started thinking more broadly about using, for instance, text as the data, rather than people. So, expanding horizons.'

For early-career researchers, interdisciplinary approaches may have positive and negative implications. Since they first need to gain sufficient expertise in their own disciplines, being part of an interdisciplinary project may pose a career risk. Outcomes and publication in larger projects may also take more time to achieve. More positively, involvement in an interdisciplinary project teaches researchers to be aware that big research problems cannot be solved within one discipline. It can also

help them to develop skills in communicating with different people and to respect other disciplines, while sparking new ideas that may contribute to future collaborations. As noted by one team member: 'I learned a lot from others in the project. … that reflects two things: firstly, we have made considerable intellectual headway in our thinking about the area, we're already breaking into new ground, thinking about new projects; and secondly, we work so well together.'

## Engagement and Dissemination

The BIAS project aims to engage with a number of stakeholders through ongoing discussion, exchange, and knowledge sharing. Key stakeholders include public and private sector employers using AI for human resource management; job seekers and employees' encountering AI in their working lives; developers, computer scientists and technicians involved in the design, testing and maintenance of algorithms and job platforms; organisations regulating the use of AI; academic and policy researchers working on related issues; and the general public who are interested in and undergoing digital transformations in society.

In the initial stages of the project, stakeholder engagement was intentionally limited to allow for a strong internal focus while team members laid the necessary conceptual and methodological groundwork for the project. In 2020, soon after the project was launched, all working relationships moved to a remote format. The team communicated entirely via digital tools (email, Zoom, Google docs) both internally, and with the organisations that they were studying. Valuable engagement took place with the advisory board, both on an individual basis with team members reaching out for advice or suggestions, and through a first annual meeting in 2020, which was held online. During the event, the team members presented preliminary results and work-in-progress, and received helpful feedback from the advisory board. A similar event was planned for June 2022.

Engagement with other stakeholders involved working with recruitment companies and employers in both countries. Data collection and engagement with job seekers was scheduled for the final year of the project, together with conference presentations, publications, and other knowledge dissemination activities to enable active engagement with academic and policy researchers, developers and computer scientists, and organisations involved in regulating the use of AI.

Decisions about publications and other intellectual outputs from the project also reflected the team's philosophy, involving discussion of what would be of most benefit to individual team members. In the internal survey, all team members rated the anticipated publication outcomes highly, suggesting a positive evaluation of the team's egalitarian approach. In some specialist fields, publications in monodisciplinary journals are the primary outlets recognised by institutions, and interdisciplinary journals are considered to be less prestigious. While there are exceptions, such as the journal *Nature*, such outlets are not an obvious choice for management, sociology or computer science scholars. The

team accepted that targeting such 'higher risk, higher return' journals would require prolonged investment, but could also result in a more significant intellectual contribution to interdisciplinarity.

## LESSONS LEARNT

By reflecting upon the objectives of BIAS project, the strategies developed for overcoming challenges, and the emerging implications at an early stage in the research process, the project team identified some concrete 'lessons learnt' that could be applied more widely to the management of international multidisciplinary research projects.

- Interdisciplinary project teams are more easily formed when there are pre-existing professional relationships developed through conferences, professional events, and working at the same institution.
- Research institutions and funders can enhance the capacity for interdisciplinary research through structural supports that connect researchers across different disciplines, faculties and across countries: for example by including joint multidisciplinary training events, within or between universities, as well as the formation of research centres and research theme groups involving academics who are working on similar (broad) issues but in different disciplines and cultural settings.
- Effective teams require a diverse mixture of skills, disciplinary orientations, and training, as well as an 'interdisciplinary mindset' and an openness to learning.
- Specific tasks should be distributed among team members with appropriate skill sets to ensure that their contributions fit into the big picture and match the original grant proposal.
- Reaching agreement in interdisciplinary projects poses challenges that can be resolved by dividing work tasks across smaller intra-disciplinary units, and allowing collaboration with and between smaller units and the larger team.
- Intellectual and administrative coordination through a 'super bridger' can help to ensure the integration of results, to enable small intra-disciplinary teams to stay informed about each other's findings, to generate pathways for additional collaborations, and provide valuable project management support for the principal investigators.
- Interdisciplinary research can bolster methodological development through rigour and the innovative use of research methodologies, and visual tools can help to 'translate' across disciplinary vocabularies and avoid potential misunderstandings.
- An egalitarian 'flat' approach to the division of labour can be a successful collaboration strategy, provided that an equitable balance is struck between empowering the project members to improve and innovate at their own pace and ensure the process remains manageable.

- Openness and flexibility are important in establishing acceptance and dialogue between researchers belonging to different disciplines, particularly when deciding on intellectual outputs from the project.
- In international projects that involve researchers based in different countries, remote meetings with synchronous communication (Teams, Zoom, Skype) offer an effective means of maintaining 'visibility' and a high level of engagement among the team members, while also preventing research-related misunderstandings and roadblocks.
- Regular communication within and across sub-projects or work packages can include monthly whole-team multidisciplinary meetings, as well as bi-weekly intra- and inter-disciplinary meetings of smaller teams working on similar tasks.
- Both planned and unplanned communication are important in keeping immediate tasks and broader project goals on track, and in responding to unforeseen issues that are common in interdisciplinary work that is focussed on evolving concepts and methods.
- Frank and open discussions need to take place at an early stage about what different team members want to get out of the project, including the negotiation of preferences associated with project disciplines for different research outputs.
- Project managers need to remain attentive to the training needs of early career researchers associated with the project, and to assist them in building a publication portfolio.

## REFERENCES


Fiore, S. M., Hoffman, R. R., & Salas, E. (2008). Learning and performance across disciplines: An epilogue for moving multidisciplinary research toward an interdisciplinary science of expertise. *Military Psychology*, *20*(sup1), S155-S170. DOI: 10.1080/08995600701804939

Hu, S., Alshehabi Al-Ani, J., Hughes, K. D., Denier, N., Konnikov, A., Ding, L., Xie, J., Hu, Y., Tarafdar, M., Jiang, B., Kong, L., & Dai, H. (2022). Balancing gender bias in job advertisements with text-level bias mitigation. *Frontiers in Big Data.* DOI: 10.3389/fdata.2022.805713

Monteiro, M. & Keating, E. (2009). Managing misunderstandings: The role of language in interdisciplinary scientific collaboration. *Science Communication*, *31*(1), 6-28. DOI: 10.1177/1075547008330922

Tarafdar, M. & Davison, R. M. (2018). Research in information systems: Intra-disciplinary and inter-disciplinary approaches. *Journal of the Association for Information Systems*, *19*(6), 523-551. Available at: https://aisel.aisnet.org/jais/vol19/iss6/2