

# Written Evidence Submitted by the Oxford Internet Institute and the University of Exeter (GAI0024)

## Executive Summary

This report was produced by Dr Callum Cant, Dr Funda Ustek-Spilda, Dr Matthew Cole, and Prof. Mark Graham, all of the Oxford Internet Institute, University of Oxford, and Dr James Muldoon of the University of Exeter. It is based on the 'AI for Fair Work' project, supported by the Global Partnership on Artificial Intelligence (GPAI), which builds on the 2019 OECD 'Recommendations on Artificial Intelligence' to develop specific and measurable benchmarks for the fair deployment of AI systems in the workplace. As such, the following evidence focuses primarily on governance from the perspective of the workplace: the location where huge numbers of people will have deep and sustained interactions with AI systems for the first time.

We recommend:

- **Creating AI-specific legislation** that can act as the lynchpin of a wider regulatory approach.
  - This legislation should meet four criteria: it should build on existing human rights and scale in proportion to risk; with specific reference to the workplace, it should place clear limits on the authority of managers over workers; it should also interface between relevant existing areas of law such as data protection, anti-discrimination and occupational safety and health; and finally, it should embed social dialogue as an essential means for resolving the differing interests of workplace stakeholders and guaranteeing civil liberties in the workplace.
- **Promoting social dialogue**, both on questions of technological development through specific legislative provision and more generally through reforms including: a right to access the workplace for trade unions; sectoral collective bargaining; and worker membership on company boards.
- **Supporting the AI regulatory ecosystem** through increases in the enforcement resources afforded to both the Health and Safety Executive and the Information Commissioners Office to implement AI regulation.
- **Gathering better data on AI system deployment** in the UK to facilitate policy making.

## Background

The concept of 'Artificial Intelligence' (AI) covers a range of new technologies that can be applied to tasks ranging from natural language processing to prediction algorithms and board games. As a result, developments in AI impact a wide variety of sectors of work and reshape individuals' lives in complex ways. It has the capacity to catalyse productivity gains, but it also poses significant risks. We have already seen how AI can contribute to enhanced surveillance capacities, discriminatory algorithms and a degradation of working conditions and workers' well-being.

For the past year, we have been carrying out a project funded by the GPAI to establish concrete benchmarks for the fair implementation of AI systems in the workplace. In order to achieve this, we conducted a global tripartite consultation with key stakeholders including Uber, Microsoft, and the International Labour Organisation. This consultation consisted of three stages: first, we developed an initial draft of the principles through a literature review; then we conducted in-depth interviews with 21 key global stakeholders, followed by a multi-participant focus group; finally, we conducted a

wide-ranging survey to gather additional data on a later draft. Our final report was published on 22 November 2022 (Cant et al. 2022). The following evidence is based on the outcomes of this research.

## How effective is current governance of AI in the UK?

Until recently, the UK Government has not had a unified and coherent approach to AI governance. The Government announced its intentions to provide such an approach in its *National AI Strategy* (HM Government 2021), which proposed ‘establishing an AI governance framework that addresses the unique challenges and opportunities of AI, while being flexible, proportionate and without creating unnecessary burdens.’ This document contained promising signs that UK regulators would be empowered to set an example for the safe and ethical deployment of AI. The UK Government’s approach to AI governance has been indicated in a recent publication, *AI Governance and Regulation: Policy Statement* (HM Government 2022), which sets out a ‘pro-innovation’ approach to regulating AI.

The current governance of AI consists of a patchwork of laws and policies which govern different sectors. Many of these laws were not written to specifically address AI and are therefore not built for purpose. This includes laws for the protection of personal data, the regulation of online spaces and the operation of algorithms. Regulators such as the Information Commissioner’s Office, the Equality and Human Rights Commission and the Health and Safety Executive have begun to take action on developing consistent standards for guiding the development of AI. However, there is no single unified approach concerning how AI should be governed that attempts to mitigate potential harms, or what kind of specific technological developments would be counted in this approach. The absence of AI-specific legislation has resulted in uncertainty over how existing legislation applies to new developments in AI and inconsistencies between regulators in how their powers should be applied.

The data available to assess AI system deployment in the UK, and subsequently the success or failure of the existing governance model, is unfortunately limited. In order to improve our understanding of questions such as these, we recommend improving the available data on AI. This could be achieved by supporting the ONS to gather information on the technology as part of regular labour market reporting. Questions of interest might include: what AI systems are being deployed and at what scale; how many enterprises are using AI systems; how many worksites are AI systems being deployed in; how many workers are working alongside AI systems; how many labour disputes are taking place in which AI systems are a point of contention; what changes are occurring in wages, productivity, rates of workplace injury and employment in worksites where AI is being deployed. This would also facilitate more informed policymaking as the rollout of AI into wider society continues.

In summary:

- There are promising signs that UK regulators are being empowered to define standards of safety and ethics in AI deployment. This is a step forwards from what was, until recently, a disorganised status quo.
- But the lack of AI-specific legislation reduces the effectiveness of the emerging governance model. This lack of central reference point is likely to lead to regulatory inconsistency and confusion.
- Better national data would allow for a stronger analysis of the successes and failures of UK AI governance.

## What are the current strengths and weaknesses of current arrangements, including for research?

In its recent Policy Statement on AI, the UK Government stated that it does not plan to introduce new legislation on AI and that it favours the adoption of high-level principles to help industry create voluntary measures to guide the production and deployment of new AI. The government's current approach has several distinct strengths and weaknesses:

### Strengths:

- **Approach is context-specific:** the Government recognises that the use cases of AI are varied and complex and will require different approaches depending on the sector and the technology.
- **Relevant regulators are identified:** the Government's Policy Statement identified the Information Commissioner's Office (ICO), Competition and Markets Authority (CMA), Ofcom, Medicine and Healthcare Regulatory Authority (MHRA), and Equality and Human Rights Commission (EHRC) as the most important regulators for the governance of AI.
- **Establishes key principles for the regulation of AI:** The UK government wants its regulators to follow six principles when devising AI rules. Many of these principles mirror widely accepted approaches to governing AI that focus on concerns over safety, security, transparency, explainability, fairness and contestability.

### Weaknesses:

- **Reliance on voluntary measures and self-regulation:** The risk of voluntary measures is that industry adopts vague and overly permissive principles that fail to guard against the unethical and dangerous deployment of AI, as they are left to interpret the principles and guidance. This risk is immediate and considerable given the recent history of the tech sector, which has consistently demonstrated an inability to effectively self-regulate. Data breaches, unlawful use of customers' data, firing whistle blowers and the suppression of unfavourable internal research have all recently occurred in many of the largest tech companies that are now responsible for developing AI. The introduction of new legislation on AI that creates an enforceable and consistent framework would limit the potentially harmful development of AI.
- **Appetite for risk is unacceptably high:** The UK Government explicitly defines their strategy as a 'risk based' approach to governing AI in which industry is left to self-regulate the majority of their practices. This strategy underestimates the potential negative consequences of poorly-designed and nefarious AI and overlooks the uncertain implications of majority of AI development. The development of advanced AI capacities has the potential to irreversibly change the structure and direction of human societies. While we are still in the early stages of its development, the next decades will create new path dependencies and value lock-ins that will have significant impacts on almost every aspect of individuals' lives. A consistent legal framework would better enable our values to be codified in new AI at every stage of the AI lifecycle and would bring the much-needed standardisation across sectors and industries
- **Lack of worker voice:** Workers should be able to participate in the development and deployment of AI, particularly when it directly relates to the management of the labour process. There are currently large asymmetries in power between stakeholders in the deployment of AI in the workplace which lead to workers' perspective being widely

neglected and deprioritised. We advise that a framework be established that allows for the codetermination of the implementation of AI in the workplace through a multistakeholder initiative that facilitates collective bargaining between workers and management.

- **Potential inconsistent application of regulatory principles:** the UK's light-touch approach and a lack of a unified legal framework is designed to increase innovation, but it could have the opposite effect. A clear and consistent set of laws governing the use of AI could help foster economic growth and innovation by providing appropriate guidance to industry. In the absence of a clear approach, UK businesses will need to second guess which complex set of laws and policies apply to their potential new developments. This will disadvantage small and medium enterprises that do not have large budgets for legal expenses.
- **No mention of the sustainable development of AI:** The training of certain AI models requires the energy-intensive use of computational power which can have significant consequences for the long-term sustainability of such projects. Principles developed for the governance of AI should pay attention to issues of environmental justice and the increased greenhouse gas emissions created by research in AI.
- **Not all AI is created equal.** Technologies are not neutral, and AI is no different. AI systems that are specifically created with profit in mind may not always work well with industries or services that are traditionally considered non-profit, or those that may not have a direct possibility of profitability. Principles developed for the governance of AI should pay attention to issues of social justice, and evaluate whether AI provides relevant and appropriate solutions.

## What measures could make the use of AI more transparent and explainable to the public?

Machine Learning (ML) is a fundamental precondition of many contemporary AI systems. ML facilitates a non-symbolic approach to AI that is not based on a human concepts and/or logic but rather the application of complex statistical methods to very large datasets. The weights attributed to any particular node of a convolution neural network, for instance, do not 'represent' a thought or idea, but rather are theta ( $\theta$ ) values determined (and sometimes redetermined) by the learning process. As a result, explanations of such systems cannot present the chain of reasoning embedded within the system, but rather the weights of different nodes. In some cases, it may be possible to attribute specific weights to specific factors, but in others this may prove more difficult. We leave the discussion of the exact nature of these barriers and the possible solutions to technical experts (See for example Adadi and Berrada 2018; Angelov and Soares 2019; Angelov et al. 2021; Dosilovic, Brcic, and Hlupic 2018). However, it is clear that ML presents partial technical obstacles to the goal of total explainability.

However, these barriers cannot be allowed to delimit the scope of regulation for two reasons: first, the necessity of explainability is so significant that it should act as a regulatory red line, which is to say that any system that cannot meet it should not be deployed (see Cant et al. 2022); and second, because regulation will incentivise actors to devote resources to meeting the technical challenge of explanation, thereby shaping technological development in a more transparent direction. So, we believe that explainability can best be summarised as both a challenge and a necessity. Governance approaches should build on this approach to demand high levels of explainability from those actors deploying AI systems, particularly in high-risk environments like the workplace.

In order to identify what kind of measures can provide explanations, it is first important to identify what *kind* of explanations we are interested in. The GPAI principles specify two kinds of explanation

owed by organisations to their workers.<sup>1</sup> The first is a ‘model behaviour’ level explanation (Samek and Müller 2019), which lays out the general processes that guide the decision making process. This should include, for instance, detail on: data features; system outcomes; and the points of interaction between the system and the labour process. These explanations could be produced by any one of a number of suitable methods. The second is the specific explanation of decisions. Rather than focusing on model behaviour, such explanations should offer workers impacted by a decision a rationale for an outcome. This overlaps with the vital and widely-agreed principle of accountability and the subsequent right to appeal decisions made by in part automated means, which we also cover in detail in the GPAI report.

These dual rights to explanation must be situated within a framework that supports the comprehension of explanations (Gilpin et al. 2019). It is not enough for employers to present workers with giant datasets and raw code and expect them to work out the details. The right to an explanation must be made into a positive right by the actor introducing the AI system through the creation of systems of support, training, guidance and inquiry that ensure that explanations are comprehensible. In circumstances where workers are represented by a trade union, the employer must also extend this offer to representatives of the union.

In summary:

- Technical barriers to explainability should not determine the scope of regulatory frameworks. Instead, explainability should be an essential precondition that actors must meet before deploying AI systems.
- Workers specifically should have a right to two kinds of explanations: both at the level of ‘model behaviour’ and the level of specific decisions.
- These dual rights must be supported by a framework that supports comprehension, including, for instance, training and support. The burden of providing this should be shouldered by the actor introducing the AI system.

## How should the use of AI be regulated, and which body or bodies should provide regulatory oversight?

A range of agencies should have areas of defined responsibility for both enforcement and prevention of AI regulation, depending on the context in which AI systems are deployed. As a workplace technology with its roots in the use of large datasets, AI is of great significance to both labour inspectorates and information regulators. In the UK, this means the Information Commissioners office and the Health and Safety Executive both have a vital role to play.

Labour inspection systems are a fundamental part of effective labour regulation across the world (Arrigo, Casale, and Fasani 2011). However, effective labour inspection relies on a suitable institutional framework and sufficient resourcing. The rollout of AI systems demands that the HSE play a variety of roles, including: ensuring updated labour regulation is being implemented via inspections and sanctions; externally scrutinising audits and impact assessments; and maintaining Occupational Safety and Health (OSH) by investigating workplaces with excessive levels of issues associated with work intensification, such as musculoskeletal injuries and stress. However, we are

---

<sup>1</sup> The definition of ‘workers’ used in the report is a broad one, which does not specify the exact nature of employment, self-employment, or limb (b) worker status. For more on our approach to the employment relation in the context of recent developments in the platform economy, see ‘Protecting Workers in the UK Platform Economy’ (Howson et al. 2021).

concerned that research into the state of UK OSH enforcement current concludes that the credibility of the British system of OSH self-regulation is at risk due to 58% real terms reductions in HSE funding between 2009-10 and 2019-20, which contributed to a 72% decline in inspections over the same period (Moretta, Tombs, and Whyte 2022). This suggests that without significant additional resourcing, the HSE will not be able to adequately monitor AI system deployment in a way which guarantees the safety of workers.

Information regulators also have a vital role to play in the governance of AI as a result of the centrality of data to ML-based approaches. Similarly to labour inspection, the effective functioning of information regulators like the ICO relies on both institutional support and appropriate resourcing. We know that ICO enforcement is an effective tool in producing greater data security (Koutroumpis, Ravasan, and Tarannum 2022), but qualitative data collection suggests that ICO enforcement teams already have to prioritise cases in a way which means many breaches of relevant legislation are not seriously investigated (Ceross 2018). As a result, we have parallel concerns to those expressed above regarding the level of enforcement which underpins the UK's overall philosophy of preventative self-regulation on questions of data.

As a result, our overriding recommendation is that as the roles of the HSE and the ICO change and expand in response to AI system deployment, these changes are accompanied by an appropriate increase in resourcing. If there is no enforcement capacity underlying the UK's approach to AI governance, then compliance is likely to be highly uneven. Those on the wrong end of social power dynamics would bear the brunt of the risks resulting from this unevenness, whilst receiving little of the reward.

In summary:

- The HSE and ICO will have a vital role to play in enforcing regulation regarding the use of AI in the workplace in any effective governance system.
- However, we have serious concerns about the under-resourcing of both regulators. Regulatory systems that deal with risks as severe as those posed by AI must be built on a foundation of solid enforcement.

## Is more legislation or better guidance required?

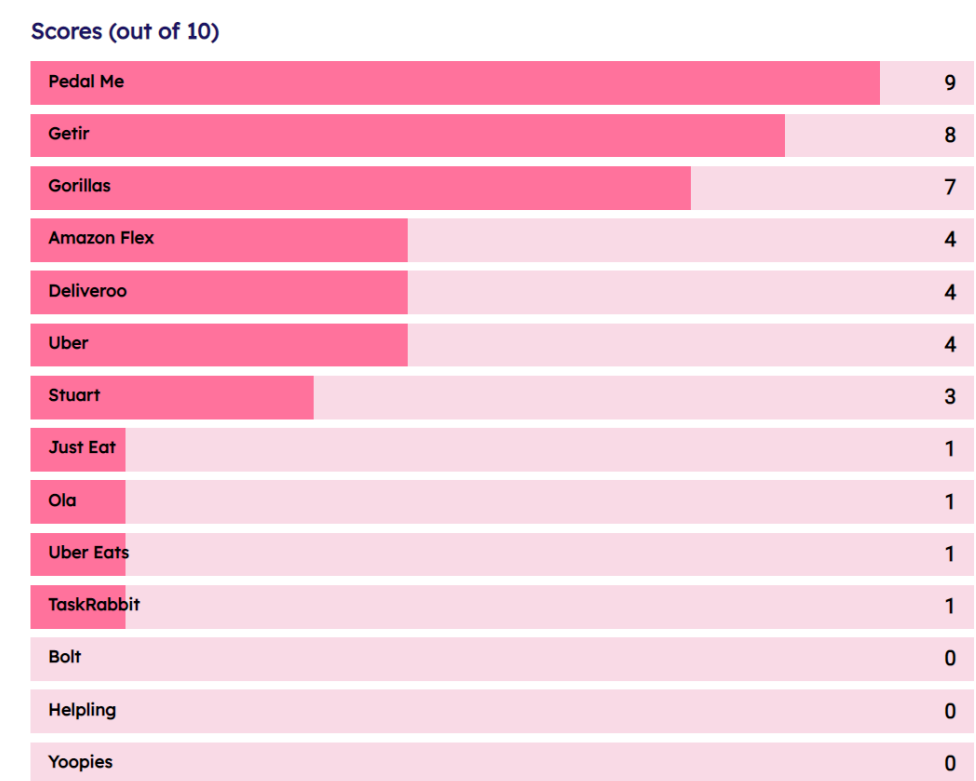
AI regulation should be based on specific legislation that builds on existing human rights and scales in proportion to risk (Mantelero 2019). With specific reference to the workplace, it should place clear limits on the authority of managers over workers (De Stefano 2018). It should also interface between relevant existing areas of law such as data protection, anti-discrimination and occupational safety and health (Aloisi and De Stefano 2022). And finally, it should embed social dialogue as an essential means for resolving the differing interests of workplace stakeholders and guaranteeing civil liberties in the workplace (De Stefano 2020). This legislation should act as central lynchpin of the regulatory system. AI systems in the workplace should be considered high risk, in line with the draft EU AI act, and should be subject to the strongest regulation.

But legislation alone is not a panacea. Guidance has a role to play in providing examples of best practice and helping shape the practice of leading actors in the AI space. Principles and benchmarks such as those published by the GPAI can help shape the debate around how legislation should be formed and implemented (Cant et al. 2022). However, as reporting by Rodrigo Ochigame has made very clear, the drive to replace binding regulation with voluntary guidance had often been part of an anti-legislative strategy pursued by powerful private sector actors (Ochigame 2019). This has led to widespread concerns about 'ethics washing' – or the adoption of voluntary ethics codes that are

promoted as an alternative to regulation but which end up being overly-permissive and failing to actually mitigate the risks of AI systems (Bietti 2021; Hagendorff 2020; Floridi 2019).

In relation to another novel technology currently having a major impact on global labour markets, digital labour platforms, the Oxford-based Fairwork research project has gathered extensive evidence on fairness and the necessity of legislation. Fairwork scores platforms in 38 countries against a set of 5 principles of Fair Work, developed through extensive international consultation. A platform that score ten out of ten meets all of our basic minimum standards for fairness. But most digital platforms fall far short of this. See, for instance, the 2022 scores for the UK:

Figure 1. UK Fairwork scores, 2022



This finding has been repeated across a range of national contexts, across both high and low income countries. Gradual progress is being made in meeting the Fairwork standards by some platforms, but self-regulation is not proving effective despite widespread notional commitments to fairness being shared amongst many major platforms. That is why we believe new regulation is necessary to improve fairness across the board. This belief is supported by the experiences of Fairwork researchers in Chile, Ecuador and Peru, who have all been involved in the drafting of effective regulatory frameworks that promote fairness in their national platform economies. In the UK specifically, this finding has motivated our policy support for responsive regulation (Howson et al. 2021). Given the incentive structure of our economy, actors who are left to self-regulate tend to engage in a race to the bottom rather than self-regulating to a point of fair equilibrium.

Drawing on these two lessons, then, we can conclude that regulation based on robust legislation that is combined with elements of voluntary guidance offers the strongest way forward for AI governance. We suggest that the UK's current governance framework will not be adequate to manage the risks of AI system development, because it lacks central legislation and developed enforcement capacities to go with it. The OECD has said that AI systems should be fair, equal, and

socially just (OECD 2019). This end point cannot, at the moment, be guaranteed without further action.

In summary:

- AI regulation should be based on specific legislation that meets four criteria: it should build on existing human rights and scale in proportion to risk; with specific reference to the workplace, it should place clear limits on the authority of managers over workers; it should also interface between relevant existing areas of law such as data protection, anti-discrimination and occupational safety and health; and finally, it should embed social dialogue as an essential means for resolving the differing interests of workplace stakeholders and guaranteeing civil liberties in the workplace.
- Guidance does have a role to play in establishing best practice. However, there is a significant risk of 'ethics washing,' and self-regulation through voluntary guidance will not be sufficient.

## What lessons, if any, can the UK learn from other countries on AI governance?

Successful AI governance will show significant overlaps with successful technological governance more generally. The novelty of AI technologies are often overemphasised, in part due to the incentive structures generated by venture capital funding and market competition, and this can lead to policymakers and others failing to see important continuities between information technology more generally and AI in particular (Cole et al. 2022; Kaltheuner 2021). So, we argue that it is both possible and necessary to draw lessons on AI governance from non-AI experiences.

Varieties of Capitalism researchers have argued that Coordinated Market Economies (CMEs) give rise to more incremental approaches to technological innovation, whereas Liberal Market Economies (LMEs) are characterised by more radical innovation patterns (Hall and Soskice 2001). When the same broad job role is compared across CMEs and LMEs, workers in CMEs tend to experience better outcomes when evaluated in terms of job quality (Esser and Olsen 2012; Frege and Godard 2014; Holman et al. 2009). Our research suggests that these divergent approaches to technological innovation are evident in AI system deployment, and that incremental approaches may still be providing workers with better job quality. When questioned about the possible danger of job destruction associated with AI and the distribution of productivity gains achieved by AI system deployment, a representative of the International Transport Federation provided us with the example of a warehouse in Norway where a combination of local institutional interaction with the labour market and social dialogue allowed for a positive introduction of AI systems:

[Often, employers argue that] “[the existing workforce] can't learn new things and therefore we're going to have to get rid of them and bring in graduates.” And it's just not true. The case that I'm thinking of is a warehouse in Norway where they retrained the existing warehouse workers and truck drivers, and they're running a highly automated warehouse now doing basic programming and, you know, sensor maintenance and all sorts of stuff.

They identified how when the introduction of AI systems is combined with social dialogue, the level of worker engagement in the process tends to be higher:

I think workers tend to be more engaged and feel ownership over technology in places where they've been part of the design. In other words, part of defining the problem to be solved and part of designing the solution to the problem. And we've seen this, for example,



in a warehouse in Norway where the where the workers talk about “my machines. [They say] “I love my machines.” You know, that way they're very positive about the impacts of technology [but when] it happens without them it usually brings about resistance [...] Part of the dignity of work is understanding what you do and how it fits in with a with a greater process. One of the frustrations that workers have with digitalization and with the process of automation in particular is that it reduces them from understanding the system to understanding the part.

This leads us to a provisional conclusion that a move towards incremental technological innovation facilitated through social dialogue in the UK would be likely to produce better outcomes for the workers impacted. Broader collective bargaining coverage and well-enforced regulation would both do much to improve outcomes. This conclusion is in line with both the existing literature on technology and job quality in high-income countries, and the data we have gathered specifically on AI system deployment, and contributed to our decision to embed collective worker voice in the GPAI principles (Cant et al. 2022). Our report adds to a growing body of research (for example, De Stefano and Taes 2021) which argues that processes of multi-stakeholder negotiation are essential to maximise the benefits and minimise the risks of AI system deployment. Governments have a clear role to play in facilitating this by creating an environment conducive to social dialogue.

We recommend:

- AI regulation should embed direct provisions for enterprise and sector-level collective technology agreements in law, and specifically prohibit any use of workplace surveillance technology to hinder the right to organise. Legislative examples such as the 1948 Italian constitution draw on social dialogue as a method to defend civil liberties in the workplace. It is the ability of collective bargaining mechanisms to consider specific contexts and negotiate the interests of multiple stakeholders that makes them highly effective in translating ethical principle into practice (De Stefano and Taes 2021; De Stefano 2020).
- The UK industrial relations system should be reviewed with the goal of introducing reforms that support social dialogue in general, such as introducing: a right to access the workplace for trade unions; sectoral collective bargaining; and worker membership on company boards.

However, this approach seems to be in contradiction with the current direction of travel in the UK. Labour lawyers have expressed concern at the increase in ‘authoritarian’ trade union legislation in the last decade (Bogg 2016) and there has been further discussion of legislation which continues this direction of travel. Such a development would have highly negative impacts on AI governance by further reducing the role of collective workers representatives in workplace decision making. Instead, the government should be looking to facilitate processes of participatory workplace redesign if AI deployment is to offer workers more benefits than risks.

In summary:

- Successful AI governance does not need to be invented out of thin air – existing models can be adapted.
- Social dialogue is a vital mechanism for the translation of ethical principles into ethical practices, and should be fundamentally embedded into the governance of AI in the UK.

- We recommend that specific provision is made for collectively bargained technology agreements in relevant legislation, and that more general changes are made to the UK industrial relations system to promote social dialogue

**(November 2022)**

## References

- Adadi, Amina, and Mohammed Berrada. 2018. "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)." *IEEE Access* 6: 52138–60. <https://doi.org/10.1109/ACCESS.2018.2870052>.
- Aloisi, Antonio, and Valerio De Stefano. 2022. *Your Boss Is an Algorithm: Artificial Intelligence, Platform Work and Labour*. Hart Publishing. <https://doi.org/10.5040/9781509953219>.
- Angelov, Plamen, and Eduardo Soares. 2019. "Towards Explainable Deep Neural Networks (XDNN)." *ArXiv:1912.02523 [Cs]*, December. <http://arxiv.org/abs/1912.02523>.
- Angelov, Plamen, Eduardo A. Soares, Richard Jiang, Nicholas I. Arnold, and Peter M. Atkinson. 2021. "Explainable Artificial Intelligence: An Analytical Review." *WIREs Data Mining and Knowledge Discovery* 11 (5). <https://doi.org/10.1002/widm.1424>.
- Arrigo, Gianni, Giuseppe Casale, and Mario Fasani. 2011. *A Guide to Selected Labour Inspection Systems: (With Special Reference to OSH)*. Geneva: ILO.
- Bietti, Elettra. 2021. "From Ethics Washing to Ethics Bashing: A Moral Philosophy View on Tech Ethics." *Journal of Social Computing* 2 (3): 266–83. <https://doi.org/10.23919/JSC.2021.0031>.
- Bogg, Alan. 2016. "Beyond Neo-Liberalism: The Trade Union Act 2016 and the Authoritarian State." *Industrial Law Journal* 45 (3): 299–336.
- Cant, Callum, Matthew Cole, Funda Ustek Spilda, and Mark Graham. 2022. "AI for Fair Work." Paris: Global Partnership on Artificial Intelligence. <https://www.gpai.ai/projects/future-of-work/AI-for-fair-work-report.pdf>.
- Ceross, Aaron. 2018. "Examining Data Protection Enforcement Actions through Qualitative Interviews and Data Exploration." *International Review of Law, Computers & Technology* 32 (1): 99–117. <https://doi.org/10.1080/13600869.2018.1418143>.
- Cole, Matthew, Callum Cant, Funda Ustek Spilda, and Mark Graham. 2022. "Politics by Automatic Means? A Critique of Artificial Intelligence Ethics at Work." *Frontiers in Artificial Intelligence* 5 (July): 869114. <https://doi.org/10.3389/frai.2022.869114>.
- De Stefano, Valerio. 2018. "'Negotiating the Algorithm': Automation, Artificial Intelligence and Labour Protection." 246. Working Paper. Geneva: International Labour Office. [https://www.ilo.org/wcmsp5/groups/public/---ed\\_emp/---emp\\_policy/documents/publication/wcms\\_634157.pdf](https://www.ilo.org/wcmsp5/groups/public/---ed_emp/---emp_policy/documents/publication/wcms_634157.pdf).
- . 2020. "'Masters and Servers': Collective Labour Rights and Private Government in the Contemporary World of Work." *International Journal of Comparative Labour Law and Industrial Relations* 36 (Issue 4): 425–44. <https://doi.org/10.54648/IJCL2020022>.
- De Stefano, Valerio, and Simon Taes. 2021. "Algorithmic Management and Collective Bargaining." Brussels: European Trade Union Institute. <https://www.etui.org/publications/algorithmic-management-and-collective-bargaining>.
- Dosilovic, Filip Karlo, Mario Brcic, and Nikica Hlupic. 2018. "Explainable Artificial Intelligence: A Survey." In *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 0210–15. Opatija: IEEE. <https://doi.org/10.23919/MIPRO.2018.8400040>.
- Esser, I., and K. M. Olsen. 2012. "Perceived Job Quality: Autonomy and Job Security within a Multi-Level Framework." *European Sociological Review* 28 (4): 443–54. <https://doi.org/10.1093/esr/jcr009>.

- Ferragina, Emanuele, and Federico Danilo Filetti. 2022. "Labour Market Protection across Space and Time: A Revised Typology and a Taxonomy of Countries' Trajectories of Change." *Journal of European Social Policy* 32 (2): 148–65. <https://doi.org/10.1177/09589287211056222>.
- Floridi, Luciano. 2019. "Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical." *Philosophy & Technology* 32 (2): 185–93. <https://doi.org/10.1007/s13347-019-00354-x>.
- Frege, Carola, and John Godard. 2014. "Varieties of Capitalism and Job Quality: The Attainment of Civic Principles at Work in the United States and Germany." *American Sociological Review* 79 (5): 942–65. <https://doi.org/10.1177/0003122414548194>.
- Gilpin, Leilani H., Cecilia Testart, Nathaniel Fruchter, and Julius Adebayo. 2019. "Explaining Explanations to Society." <https://doi.org/10.48550/ARXIV.1901.06560>.
- Hagendorff, Thilo. 2020. "The Ethics of AI Ethics: An Evaluation of Guidelines." *Minds and Machines* 30 (1): 99–120. <https://doi.org/10.1007/s11023-020-09517-8>.
- Hall, Peter A., and David W. Soskice, eds. 2001. *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. Oxford [England] ; New York: Oxford University Press.
- HM Government. 2021. "National AI Strategy." Office for Artificial Intelligence. <https://www.gov.uk/government/publications/national-ai-strategy>.
- . 2022. "Establishing a Pro-Innovation Approach to Regulating AI." London: HM Government. <https://www.gov.uk/government/publications/establishing-a-pro-innovation-approach-to-regulating-ai/establishing-a-pro-innovation-approach-to-regulating-ai-policy-statement>.
- Holman, David, Stephen Frenkel, Ole Sørensen, and Stephen Wood. 2009. "Work Design Variation and Outcomes in Call Centers: Strategic Choice and Institutional Explanations." *ILR Review* 62 (4): 510–32. <https://doi.org/10.1177/001979390906200403>.
- Howson, Kelle, Callum Cant, Alessio Bertolini, Matthew Cole, and Mark Graham. 2021. "Protecting Workers in the UK Platform Economy." Oxford: Fairwork. <https://fair.work/en/fw/publications/13335/>.
- Kaltheuner, Frederike. 2021. *Fake AI*. <https://fakeaibook.com/>.
- Koutroumpis, Pantelis, Farshad Ravasan, and Taheya Tarannum. 2022. "(Under) Investment in Cyber Skills and Data Protection Enforcement: Evidence from Activity Logs of the UK Information Commissioner's Office." *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4179601>.
- Mantelero, Alessandro. 2019. "Artificial Intelligence and Data Protection: Challenges and Possible Remedies." Strasbourg: Consultative Committee Of The Convention For The Protection Of Individuals With Regard To Automatic Processing Of Personal Data. <https://rm.coe.int/artificial-intelligence-and-data-protection-challenges-and-possible-re/168091f8a6>.
- Moretta, Andrew, Steve Tombs, and David Whyte. 2022. "The Escalating Crisis of Health and Safety Law Enforcement in Great Britain: What Does Brexit Mean?" *International Journal of Environmental Research and Public Health* 19 (5): 3134. <https://doi.org/10.3390/ijerph19053134>.
- Ochigame, Rodrigo. 2019. "How Big Tech Manipulates Academia to Avoid Regulation." *The Intercept*. 2019. <https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/>.
- OECD. 2019. "Recommendation of the Council on Artificial Intelligence."
- Samek, Wojciech, and Klaus-Robert Müller. 2019. "Towards Explainable Artificial Intelligence." In *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*, edited by Wojciech Samek, Grégoire Montavon, Andrea Vedaldi, Lars Kai Hansen, and Klaus-Robert Müller, 11700:5–22. Lecture Notes in Computer Science. Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-030-28954-6\\_1](https://doi.org/10.1007/978-3-030-28954-6_1).