



Institute for the  
Future of Work

## Report

---

# IFOW Sandbox *Phase 1 Methodology Report*

Abby Gilbert, Joseph Marriott and Mia Leslie  
April 2026

---



*Produced with funding from*



**Trust  
for London**

# Contents

---

Executive Summary	03
.....	.....
1. Rationale	04
.....	.....
2. Design Principles for the IFOW Sandbox	06
.....	.....
3. Process	13
.....	.....
4. Methodology	17
.....	.....
5. Disclosures: Funding	21
.....	.....
6. Conclusions	22
.....	.....
References	23

---

## Citation

Gilbert, A., Marriott, J., Leslie, M., *IFOW Sandbox - Phase 1 Methodology Report*. London: Institute for the Future of Work. DOI: 10.5281/zenodo.18480973

## Permission to share

This document is published under the Creative Commons Attribution Non Commercial No Derivatives 4.0 International Licence. This allows anyone to download, reuse, reprint, distribute, and/or copy IFOW publications without written permission subject to the conditions set out in the Creative Commons Licence.

For commercial use, please contact [team@ifow.org](mailto:team@ifow.org)

---

# Executive Summary

---

In the last few years, sandboxing has moved from being a peripheral, barely mentioned concept, to a mainstay of regulatory strategy. Yet this approach is being used across highly divergent jurisdictions and regulatory strategies: in the EU, where prescriptive AI legislation has been introduced,<sup>1</sup> in America, to drive market-led deregulation,<sup>2</sup> and in the UK, where there is interest in understanding the law more holistically.<sup>3</sup>

This begs the question: what is a sandbox, and what ends should a sandbox serve? We seek to address this question in this paper via lessons we learned from the work of the IFOW Sandbox - the first (to our knowledge) established to explore the intersection of algorithmic technologies and work. We articulate the guiding design principles, processes and methodological approach taken to operating the Sandbox in its first phase, and reflect on the value of this approach within the wider ecosystem.

This has proven effective in addressing key research questions. However, we view the Sandbox as a dynamic environment, and anticipate changing this methodology to serve future research needs.

# 1. Rationale

---

The Institute for the Future of Work's Good Work Charter sets out ten fundamental dimensions or principles of 'Good Work' – work that has fair pay and conditions; work that promotes dignity, autonomy and equality; work where people are properly supported to develop their creativity, skills and capabilities; and work where they have and can express a sense of community.<sup>4</sup>

The dimensions of good work are interdependent and interrelated. The framework directs attention towards rights, freedoms and interests at work protected by a range of hard and soft laws and focuses attention on substantive, known impacts to access, pay, terms and conditions and quality of work, acting as a 'checklist' for known impact areas. But it also aims to help surface neglected, emerging and cumulative impacts and their relationships, exploring and 'testing' new evidence against a baseline, while inviting review and development over time.

The Good Work framework also acts as a normative framework which can bridge AI principles to practice – something widely understood to be the next frontier of 'Responsible AI'. Recent guidance and recommendations from the OECD, UNESCO and the Council of Europe support this approach, highlighting specific risks and impacts to the realisation of AI principles in the workplace, across the domains of good work.

Beyond this, recent research has shown that good work is central to the wellbeing and flourishing of individuals, communities and the country. Good work can help address the toughest socioeconomic challenges and rebuild strong, resilient communities. Research shows that access to good work confers protection against social and economic shocks, and that it can build and enables capabilities in ways which make individuals more resilient to technological disruption.<sup>5</sup> For these reasons, the framework doubles as an overarching orientation for mission-driven policy making, as well as for research.<sup>6</sup>

The emergence of new tools and technologies has – throughout history – presented opportunities and risks to good work. Technological disruption, the process by which technology is harnessed to change the way tasks, jobs and organisations are designed and operate, reflects choices that have been made through the period of design, development and deployment. These choices are themselves shaped by hard and soft regulation, ethical norms, behaviours, processes and cultures, knowledge and skills. The rules society creates to guide technological disruption are what ultimately determine the success or failure of nation states through major moments of technological change.<sup>7</sup> These rules are the means by which individual rights are protected, and how fairness can be ensured as the dividend of innovation.

As we undergo one of the most significant transformations to the economy since the first industrial revolution, the UK is characterised by significant distrust of government.<sup>8,9</sup> Against this backdrop, and as our research suggests, people do not experience rights and protections at work equally, and experience different levels of protection from technological disruption.<sup>10</sup> Further, there are significant gaps in protection when thinking of the different forms that automation can take.<sup>11</sup>

This does not just invite the development of regulation - it also invites reflection on how we can ensure that fairness is both perceived and realised as we develop the institutions which protect good work. Innovation is an opportunity to steer society towards something better.

The scale of current disruption presents a challenge not only to the capacity of government, but also to the very mechanisms of governing. Command and control techniques are under strain, as digital ecosystems become increasingly complex, services increasingly immaterial, value increasingly intangible, and power asymmetries increasingly polarised.

This is the context in which sandboxes are being called upon to serve highly diverse needs. For the concept of sandboxing to work, genuine commitment to regulatory innovation is required. Sandboxing as an extension of compliance logics will not be adequate to face current challenges. In many ways, these challenges require a fundamental reset in governance, one which understands how rules are interpreted, understood, and trade-offs are made in practice.

## 2. Design Principles for the IFOW Sandbox

Sandboxes have been understood as ‘a distinct type of regulatory instrument within the larger categories of flexible, experimentalist, and anticipatory modes of governance’.<sup>12</sup>

Alternatively, as the World Bank suggests: ‘Sandboxes are... useful where empirical evidence is needed to support policy development. They can be beneficial where regulatory requirements are unclear, missing or create barriers to entry disproportionate to the risks.’<sup>13</sup>

In this section, we describe the design principles used in the development of the IFOW Sandbox, according to our own definition of their potential.

*Sandboxing is an excellent approach to surfacing the impacts of novel technology, while also engaging in regulatory innovation and clarity: it can help where there is **legal, regulatory, or governance uncertainty, weak interoperability, or gaps**. The approach can be used **before, during, or after** legislation is enacted by focusing on sharp research questions which allow for feedback into regulatory design and processes.*

*This can be through **prototyping and evaluating** proposed legal, regulatory, or socio-technical governance interventions, and by using these activities to **surface new evidence** on interactions, interdependencies, inefficiencies and impacts of new technology on different governance approaches.*

*Sandbox assessments should consider regulation as part of a socio-technical system, taking into account the interactions between **technology, business models, and formal and informal rules**, as these are shaped by **cultural, psychological, sociological, political and economic factors**.*

As we explain further in this chapter, we abide by the following principles in operating the Sandbox:

- Independent
- Full Stack
- Multi-Scalar
- Interdisciplinarity
- Innovative
- Accountable
- Transparent
- Participatory

Abiding by these principles allows us to realise a socio-technical approach, which examines the application of existing and prototype rule regimes which seek to govern novel technologies as these are realised within a given social, political, cultural and economic context.

### Key characteristics of the IFOW Sandbox

**Goal:** the reflexive evaluation, development, and iteration of existing and potential rule regimes and their associated social, technical and operational architectures, to promote a wider access to good work through technological disruption.

**Scope:** the IFOW Sandbox is not delimited to the application and fit of a single regulatory regime, but considers how legislative and regulatory domains intersect and apply to real-world use cases. Specific regimes may be a focus for specific open calls or research questions.

**Scale:** in its first phase of operation the IFOW Sandbox has been operating only within the UK. Reflecting findings in Phase 1 and research questions arising from this, its operation and design are scoped to be cross-jurisdiction in Phase 2. There is a significant need for cross-border Sandbox activity.

**Regulatory Leeway:** the Sandbox operates without derogated authority, as we do not offer regulatory exemptions to participants, but study the application of existing regimes, while prototyping new governance frameworks, in order to provide insight to a range of regulators and legislators.

**Impact of Discovery:** insights from this work are used to inform intervention at three scales: businesses who participate and in the wider ecosystem through report publication; at the level of 'soft' regulatory bodies who structure the behaviour of those actors; and at the ecosystem level via engagement with formal regulators, and legislators.

## Independent

Sandboxing is commonly described as activities within a secure or 'controlled' environment. To achieve controlled conditions typically requires an 'ex-ante' (before use) approach, running a system on real or synthetic data, within a secure technical environment. Risks (rather than impacts) are modelled and evaluated without the system interacting with the real world. Such a controlled environment also aligns with ideas of 'experimenting' with the application of new or existing rules. Looking at systems before they are deployed 'in the wild' also ensures no real-world harm is caused.

Often, when a sandbox is operated by a regulator or public body, firms are incentivised by temporarily reduced or removed regulatory burdens, limitations on regulatory liability, or incentives. This is deemed necessary to encourage firms to reveal their operations more openly to the regulator, allowing for greater mutual understanding. Increased transparency contributes to greater learning. This relief typically applies to firms selected with robust admittance criteria on a case-by-case basis and is in effect only for a limited time.

Such an approach has merit. It allows a regulator to evaluate how a current rule regime does or does not apply, without them becoming complicit in real-world harm. Where there is derogated authority (law permitting the law to not be applied within the sandbox) this also allows the regulator to build trust with the participant, increasing levels of information sharing, by removing the requirement to enforce if a breach is discovered.

However, this approach presents challenges when looking to fully understand how technology is shaped by social, cultural, political and economic factors within the workplace, how it changes work and what role formal or informal rules play in this.

For a sandbox to examine the full downstream impacts of algorithmic systems, an ex-ante through to ex-post approach is arguably necessary.

Given the obligation on public bodies to not be complicit in harm, this perhaps necessitates independent approaches to sandboxing. Other potential benefits of operating independently could include increased recruitment of and access to information from industry participants, and therefore, greater aggregate regulatory insight and learning. Distance from regulators can reduce anxiety among participants about later penalty or culpability for what could be surfaced within the defined period of rule exemption after exit.

### Full Stack

Sandbox environments have been justified as tools to accelerate domestic AI companies, given evidence that passing through a sandbox increases investment.<sup>14</sup>

While this has its strengths for building a responsible, domestic AI ecosystem, it also has limitations when looking to build good work through sandboxing. Approaches to sandboxing which look to the compliance of single innovators, rather than the whole 'stack' or whole value chain involved in an innovation, against a regulatory regime, can fail to understand:

- How design, development and deployment choices by different actors within the value chain each contribute to impacts or outcomes;
- How dynamics between different actors contribute to outcomes;
- The ways in which incentives (material) and values (subjective) determine and are co-determined by those ecosystems.

In Phase 1 of operations, the IFOW Sandbox has focused on recruiting 'adopters': firms that are procuring novel technologies to reshape their operations. Our methodology in Phase 1 (see section 3) recruits the adopter, as a route to finding the relevant innovators and understanding their dynamic relationship.

This is important as firms adopting AI rarely design, develop and deploy these systems. The vast majority of 'adopters' purchase from third-party providers. Nearly 80% of businesses report accessing, buying, licensing or using third-party AI tools, while more than half (53%) rely exclusively on third-party AI tools.<sup>15</sup> In turn, approaches which only examine the innovation provider - as is typical of strategies which seek to use a 'controlled environment' and 'ex-ante' approach - could lose sight of the way these systems play out in practice, and so, the nature of workplace impacts. Further, understanding the relationship between these actors is critical to devising effective governance.

Through analysing procurement and contracting data with third party providers of tools, as well as any technical audit data, this approach allows sight of what happens within the workplace, and how this is shaped 'up' the supply and or value chain.

Adopting firms are not only where the next wave of training data is being sourced, but also the 'nuclear cell form' of our wider economic transformation.<sup>16</sup> Enterprises adopting are where technology comes to interact work and wider sectoral regimes in practice. In many instances, this may require attention to the role of regulation created outside of and beyond the UK. See more in our report on Post-Deployment Training Data released with this report.

### Multi-Scalar

Decisions about AI adoption are shaped by a range of factors, ranging from local through to global. As demonstrated in our first empirical report of the IFOW Sandbox, there are sharp interdependencies between what happens within a single firm, and the global geopolitical moment. Both as this acts outside of formal systems, to shape what is or feels possible, and as this relates to more established processes or frameworks, such as international law, e.g. trade regimes governing data collection, use and processing. Further, many regulatory regimes which apply within the workplace and shape adoption practices - such as professional standards, or GDPR - are considered within the context of equivalence in trade

deals.<sup>17</sup> A multi-scalar approach follows from being interdisciplinary, and cross-regulator. Recognising this, we onboarded global governance bodies to our Regulator Plus Forum.<sup>18</sup>

While these global dynamics are significant, so too are local ones. Previous IFOW work has demonstrated that firm strategies and choices in adoption – as these shape job quality, skill demand, and availability – are shaped by local economic factors and innovation systems.<sup>19</sup> With this in mind, sandbox activity should be conscious of place, below the national regulatory scale, as a factor shaping business strategy for adoption, and so too the application and interpretation of law and ethical decisions.

Recognising these themes can be reflected in (a) selection criteria which pay regard to location (sub-national) (b) approaches to the dissemination of regulatory insight which consider the multiple agencies which could be involved in forming a response to a given challenge (c) the selection of firms based on multi- or single jurisdiction.

### Interdisciplinary

A deeply interdisciplinary lens is required to understand the interactions between rules of varying formality, their interpretation, application and outcomes. This should inform both strategies for testing and experimentation, and for feedback into the regulatory ecosystem.

We therefore draw on literatures, and expertise from our network of fellows, across disciplines such as:

- **computer science**, for instance, to surface and deconstruct how algorithmic systems function and the real technical feasibility of mitigations, redesign and alternatives;
- **critical legal studies and regulation theory**, for instance, to understand how the rules apply, are structured, and sit within wider philosophies of government;
- **sociology**, for instance to analyse how power within adoption decisions is structured and reflects wider social processes;
- **psychology**, for instance, to understand how behaviours are shaped and decisions made;
- **industrial relations**, for instance, to understand how workplace dynamics shape outcomes;
- **political economy**, for instance, to understand interfirm dynamics and value chain impacts;
- **philosophy**, for instance, to unpack questions of unquantifiable impacts to dignity, autonomy;
- **ethics**, for instance, to support reflection on what is and is not acceptable within the context of alternative strategies, before they are consolidated as rule regimes.

Interpretation and decision-making about how to apply regulatory regimes is arguably as significant to the outcomes of law as the law itself. As Ewick and Silbey (1998) suggest: ‘every time a person interprets some event in terms of legal concepts or terminology - whether to applaud or to criticise, whether to appropriate or resist - legality is produced. This production may involve innovations as well as faithful replication’. Further, while the law primarily operates through systems of compliance defining minimum standards of behaviour and mechanisms of sanction where standards are breached (determining the boundary between what is prohibited and what is permitted) this does not help to inform or seek to guide what actors ought to do, or inform how decisions about trade-offs are reached.

The IFOW Sandbox is overseen by a Regulator Plus Forum, which incorporates the expertise of, and returns insights to, a cohort of formal and informal rule-makers. This ranges from those who create ‘soft’ rules by training or upholding the professional standards of workers who are also ‘accountable agents’ within firms driving adoption (professional membership

bodies and associations) through to those upholding compliance with law, i.e. regulators, and government departments responsible for the policy environment within which business strategy is shaped, and risks and trade-offs are conceived and evaluated.

### Innovative

Sandboxes can be used to evaluate the application of current law. However, they can also be sites of ‘prefiguration’, where future regimes or processes guiding design and implementation can be practised, rehearsed and enacted.

The concept of prefigurative law – practising behaviours as if shaped by a law that does not yet exist - seems inherently contradictory. However, soft governance has long been modelled by firms as ‘corporate social responsibility.’ The question is which processes, with the agency and influence of which actors, lead to the development of new rules. In ‘Towards an adventurous institutional politics – the prefigurative ‘as if’ and the reposing of what is real’, Davina Cooper suggests that institutional change can arise through ‘prefiguration’.<sup>20</sup> Drawing from Yates (2015), this is defined as when people “enact some feature of an ‘alternative world’ in the present, as though it has already been achieved”.<sup>21</sup>

Our methodology for the Sandbox in Phase 1 of operation (see Section 4) has been to implement, as an advanced case study, a framework created under grant funding from the UK Data Regulator, the Information Commissioner’s Office. While IFOW has championed this approach being supported in primary legislation as it stands, this is a model for governance which can be used on a voluntary basis.<sup>22</sup> As this process invites firms and workers to document design choices, identify risks, reflect on and mitigate those risks, and establish monitoring frameworks, this process also doubles up as an advanced case study methodology.

In turn, via implementing this for further research within the Sandbox, we’ve been able to test a future governance regime, evaluating its utility, interpretation and efficacy while also gathering insights for wider research questions.

In future years, we would equally look to innovate new regulatory processes or frameworks, to test evaluate and implement these as part of any activity.

### Accountable

The IFOW Sandbox acts as a neutral broker and intermediary for different parties invested in ensuring that technology advances a fairer future of better work. We are accountable to this mission through:

- Our terms of reference with members of the regulator plus forum;
- Our terms of reference with industrial partners;
- The outcomes we achieve, as reported in annual reports, and extent to which we service our core objective of improving governance of AI adoption within the workplace.

Core to achieving this interdisciplinarity and remaining accountable is reflexivity; the process of recognising and scrutinising the role of the researcher (and researching institution) in what is researched, and from this, what is understood or surfaced. Barbara Myerhoff and Jay Ruby (1982) suggested that ‘Reflexive, as we use it, describes the capacity of any system of signification to turn back upon itself, to make itself its own object by referring to itself: subject and object fuse’ (p. 2).

We are also accountable to our funders, and their objectives for our work. We will publish these in our annual methodology reports.

This is particularly important as IFOW engages directly with firms and regulators, as a trusted

intermediary, serving an explicit mission. This is fulfilled in alignment with the transparency principle.

## Transparent

By taking a research question-led approach, being explicit about methodology, and publishing annual reports, we hope to contribute to the development of and conversation around the application, adequacy, interoperability and development of formal and informal rules governing technology within the workplace.

Having integrity through research-led approaches to sandboxing requires being explicit about agendas and assumptions. Evaluating the relationship between good work as a legal and a normative construct, and wider AI governance, regulation, standards, and initiatives, is an active choice. In our partnership agreements (see more in 'Process'), we are explicit about our purpose: to advance good work through understanding the conditions for its realisation through technological disruption.

Some technical approaches to the assurance of AI systems assess against defined and objective standards. This can miss previously unsurfaced, unidentified impacts arising from algorithmic systems, including those for which we may not yet have the language.<sup>23</sup> This requires more open, and qualitative approaches to data analysis which allow the unknown to be surfaced and conceptualised. These deeply qualitative interpretations, where there are not existing measures, require reflexivity and transparency the most.<sup>24</sup> Reflexivity means that the researcher is conscious of their own position within the world they are examining to understand how this shapes their understanding.

We look to uphold transparency by publishing annual reports on our work and findings, being reflexive and also ensuring that partners are transparent (via our Terms of Participation).

IFOW holds an unusual set of relationships:

- **Reaching into local communities**, understanding innovation ecosystems and labour markets as they operate sub-nationally, characterised by high levels of internal inequality – quite aside from regional national inequalities;
- **Holding close relationships with Government**, facilitating dialogue and building the capacity of public bodies to understand and evaluate AI impacts;
- **Enabling richer conversations within Parliament**, through the All Party Parliamentary Group on the Future of Work;
- **Supporting regulators** to develop an evidence base about the potential application and limits of their regimes;
- **Supporting businesses** to understand how their decision-making is impacting employers;
- **Supporting workers** to understand their rights, protections and entitlements while supporting them to use their own insight and wisdom to steer the direction of innovation.

These groups can have divergent and, at times, conflicting interests. However, holding trusting relationships with each party is a prerequisite of this being effective.

While we considered information flows between Regulator Plus forum members and participants, we steered away from this in the end, allowing us greater flexibility and protection for both regulator and firm.

We therefore act as a good faith enabler that seeks to protect participants, firms and

employees, to ensure there is no risk of reprisal while also ensuring effective innovation in regulatory processes via providing ‘feedback’ into the regulatory eco-system, sourcing frontline experience with lower risks to regulators and wider scope for investigation into practice.

### Participatory

Participation is a distinct principle in the Good Work Charter. This reflects the right and expectation that a person should be involved in decision-making where their rights, freedoms or interests are affected, including as this relates to the determination and improvement of working conditions.

In the Sandbox, participation is further considered to enable the best assessment of risks and impacts, informing regulatory development. This is because the first-hand experience of people who are (or will be) interacting with an algorithmic system is required to ascertain the breadth, nature and severity of actual and likely risks and impacts.

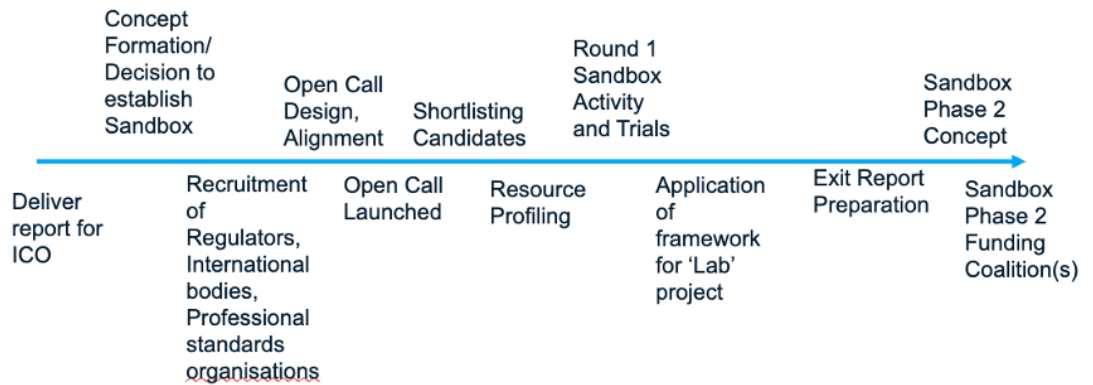
Adverse impacts and the practical or ethical concerns of workers cannot be forecast on their behalf and the impact of iterative use may change, or only be detected, over time.<sup>25</sup> This means that workers and their representatives can be seen as ‘domain experts, and need to be involved in the context-sensitive and ongoing assessment of impacts experienced across the dimensions of the Good Work Charter. This is particularly important where algorithmic systems increase information asymmetries between employers and workers.<sup>26</sup>

Worker involvement is also necessary to ascertain new, invisible or unintended consequences. Examples highlighted in international guidance on AI include impacts on people’s dignity where there is a “chilling effect” on behaviour,” and on autonomy, where workers’ ability to make decisions can become impeded. In many cases, these impacts may not have matured to the point of formal recognition as a health and occupational hazard. Here, worker involvement may unlock important information and interpretations of impacts on Good Work. In the Sandbox, we are transparent about looking to involve those who have been furthest from the conversation about AI and its better governance, frontline workers. This is an active choice to surface previously undetected risks and impacts, overcoming the risk of ‘epistemic injustice’ in AI governance.<sup>27</sup>

# 3. Process

In this section, we set out some key aspects of process in the first operating phase of the IFOW Sandbox. The following chapter will then focus on methodology.

Graphic 1: Phases of Sandbox Development to this Point



## Open Call Development

We designed the open calls to focus on use cases where there is uncertainty, weak interoperability, or gaps in formal and informal frameworks, as relating to the intersection of good work, and technology governance.

After choosing these themes, we cross-referenced with the work-plans of our partners on the Regulator Plus Forum to ensure there was no duplication of effort. This bypasses the debate about regulation being ‘outpaced’ by AI Governance.<sup>28</sup> There is already mass market penetration of non-compliant tools, owing to a combination of: pressures in the investment ecosystem which can disincentivise regulatory compliance; a ‘build fast and break things’ ideology which tends to optimise for short-term financial rewards; a lack of AI adoption skill in firms who could be more intelligent consumers’ to safeguard their own regulatory and reputational risks; and weak institutional strength for firmer approaches enforcement.<sup>29</sup>

The question is clarification and regulatory development. As IFOW is not using Sandboxing as a route to ‘market entry’ but rather as a research and innovation infrastructure to support the development of regulatory insight we believe this is justified. Further, participants remain anonymous to prevent any incidental private benefit.

## Admission Criteria

Admission criteria for sandboxes can typically cover themes such as level of public benefit, innovativeness, novelty, and level of maturity.<sup>30</sup> Transparency about terms of recruitment to Sandbox environments, particularly when run by regulators, is critical for mitigating risks relating to regulatory capture, the process by which firms gain exclusive and unwarranted advantage.<sup>31</sup>

Graphic 2: Timeline of Sandbox Activity

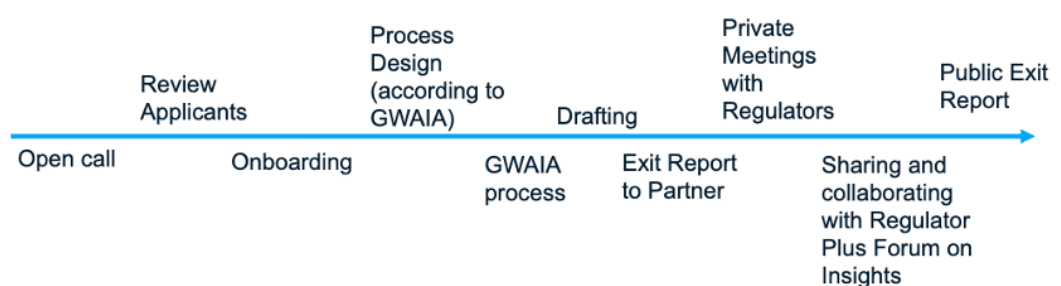


Table 1 - Open Call Use Cases and Good Work Principles

Use Case	Rationale	Good Work Principles	Core AI Governance Domains
<b>The use of Large Language Models within hiring AI</b>	The use of AI in hiring has become increasingly commonplace, with 98% of FTSE 100 companies now using it. As we have documented, a range of algorithmic functionalities that use varying sources of data and approaches to processing, can underpin these systems. Our particular focus is that of LLMs, that may be able to provide a more holistic view of a candidate's suitability for a role, compared to current keyword matching approaches used in applicant screening. The specific use of LLMs in this context presents challenges for governance regarding equality and discrimination, the effectiveness of audit tools, and prompts critical reflection on their effectiveness in the workplace.	Equality Dignity Fair Terms and Conditions	Transparency and Explainability
<b>The use of AI to model and detect methods of work</b>	Detecting work methods is key to future automation, and advancements in technologies such as Edge Computing and Digital Twins allow for real-time guidance to workers and restructure the tasks they carry out within the workplace. If done correctly, these technologies can be used to redesign jobs and their corresponding tasks, in ways that improve job quality and wellbeing. However, there are various approaches to eliciting knowledge, which may infringe on individuals' rights and freedoms. Additionally, systems have varying degrees of robustness and can change firm capabilities where procurement of SaaS solutions is not critical. This presents questions for good work, and firm sustainability.	Learning, Fair Pay Terms and Conditions	(Limits of) GDPR and Copyright
<b>The use of AI to monitor workforce physical, emotional or psychological states</b>	Advancements in computer vision and machine learning have led to the use of affective computing in the workplace to identify and interpret worker behaviours and sentiments. Whilst workplace monitoring is not a new concept, these practices have intensified post-pandemic. We aim to better understand the ways in which increases in workplace surveillance and opacity can impact fundamental rights and worker wellbeing as well as how this use of AI can generate new insights that may help support workers, and identify excessive workloads.	Dignity Wellbeing Autonomy	Health and Safety Regulation
<b>The use of AI to detect skills and competencies for promotion and work allocation</b>	The use of AI within hiring practices has been well documented. However, AI is also commonly used to allocate tasks to workers within jobs, based on perceived competencies. Further, it may do this based on different sources of data and derived analyses of the skills of individual workers. This technological innovation may allow firms to allocate jobs more effectively across their workforce, potentially increasing productivity but has associated risks. We are interested in understanding the potential impacts of the objectification of skills within the workplace, including the possibility of discrimination as well as how this technology benefits the workplace.	Equality Learning Fair Pay Terms and Conditions	Robustness Construct Validity

As IFOW does not afford access to regulators or regulatory leeway as part of the process, for us, the selection process instead focuses on:

- finding enterprise-level partners who are making a relevant adoption (as defined by their interest in one of the uses we set out in the open call)
- the enterprise being at the right stage (pre or mid- procurement) of the adoption process in order for us to follow our methodology (see Section 4).

While we would like to select for sector, size, and organisation structure, presence or absence of unions and more, however, comparative studies would depend on scale of resource.

### Recruitment process

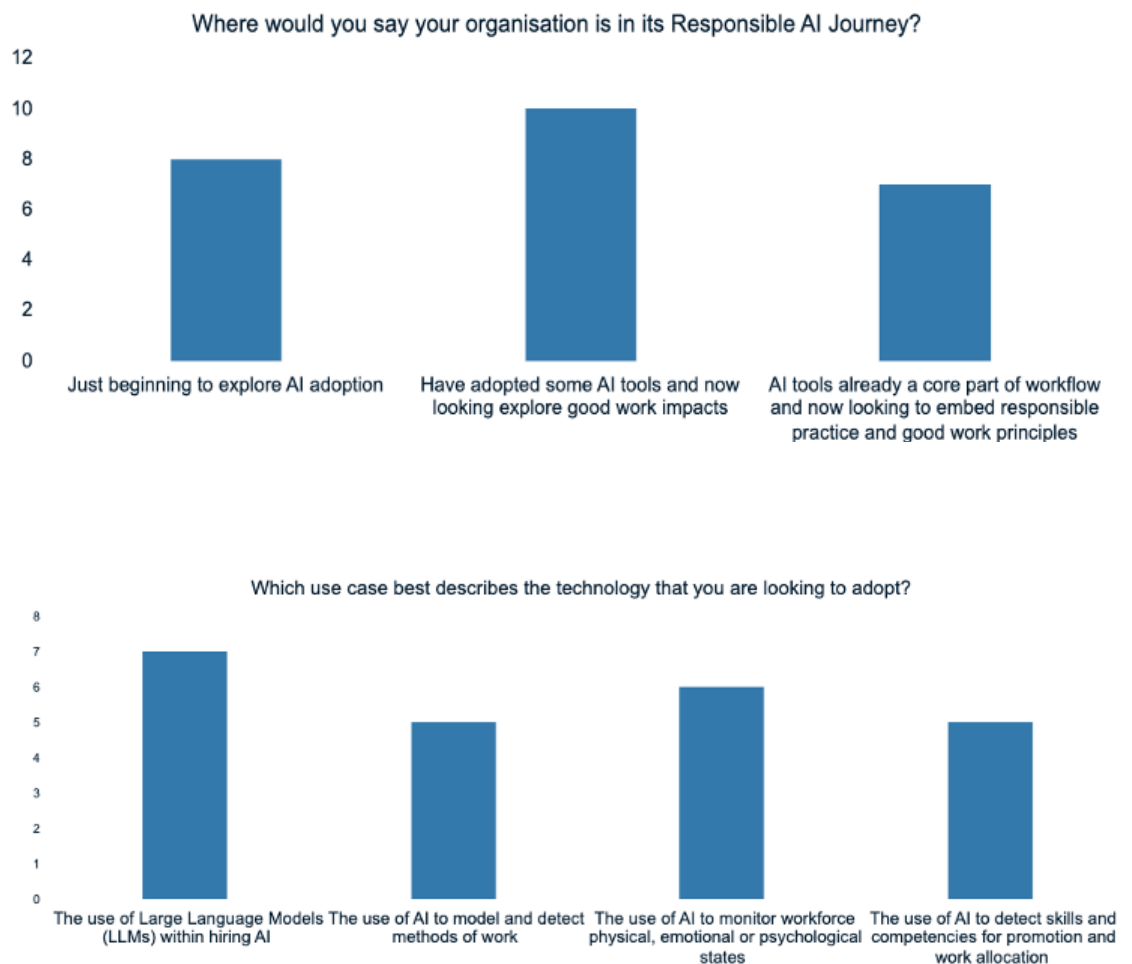
Our application form requested information on sector and state of the organisation in terms of their responsible AI journey, and the open call use case they were planning to adopt or implement. We advertised this via our website, newsletter, LinkedIn, and some industry press. A reflection was that further industry and sector press would have been useful to secure more interest and necessary if we were to attempt a comparative analysis. We technically ran two open calls. The first followed the above process. Some of the responses are below.

The latter was run with an advert which spoke to the research interests of a specific, funded project (see Funding) but in language which was recognisable by firms. This focused on recruiting businesses interested in using AI to advance productivity, as a general objective. This secured 17 Expressions of Interest. We then conducted interviews with applicants to triage them.

In some instances where it seemed aspirations were not aligned with the research process, we signposted these organisations to other available solutions, such as the DCRF Hub or ICO’s Sandbox.

The applicants were distributed across use cases, with a reasonable distribution of experience in AI Adoption.

Figure 1 - Open Call response distributions



## Terms of Participation

We set out a series of requirements from learning partner organisations via Terms of Participation. Key themes covered are presented below.

**Organisational Resource:** We asked that organisations commit the appropriate resources required to meaningfully undertake an assessment of at least one AI use case within the organisation, making clear that the time taken will depend on the complexity of the system and approach to the methodology. We set an expectation that organisations would secure buy-in from senior leadership to participate. Before participating in the Sandbox, IFOW provided bespoke timelines and resource estimates for each organisation to participate based on a co-developed Sandbox Plan. Organisations were asked to make available at least one senior member of management. In large organisations, this may be the Chief Executive Officer, Chief Operating Officer, Head/Director of People, Chief Technology Officer, Chief Information Officer, Head of Operations, or Data Protection Officer, etc. As well as a named lead, we asked that organisations commit allocation of staff time and resources to engage with the process.

**Good Work:** Maintaining our commitment to transparency and trust, we made explicit reference to our interest in advancing good work through the learning partnership. Terms of participation made clear that IFOW was not accountable for choices made by the learning partner, and that the learning partner continued to hold all accountability for compliance with applicable legal regimes. We were further clear that the guidance we provided in following our proposed methodology did not render IFOW liable for any decisions made by the learning partner or preclude them from seeking legal advice. All regulatory obligations still applied to the learning partner.

**Transparency and Confidential Information:** as part of the terms of the Sandbox, partners had an obligation to surface, create and disclose to IFOW information relating to the key choices made in the course of Good Work Algorithmic Impact Assessment (the methodology). The Participant was asked to ensure that it has all necessary rights and licenses in place to enable access to relevant information for the purposes of the Learning Partnership. This included:

- procurement tenders and contracts
- agreements between the Participant and the developer such as non-disclosure agreements (NDAs), data sharing agreements, and service level agreements
- technical auditing information. Our correspondence with Learning Partners was recognised as insight which would be used to inform potential publications. IFOW was permitted to disclose (on an anonymised basis) any non-confidential information to the extent that such disclosure is reasonable in relation to the Sandbox and that the partner did not specify that this was not appropriate.

**Termination:** In accordance with the Sandbox Plan, we set a defined Term of Participation. If the process was not completed in this period, an extension could be granted. If either partner felt there was not a sufficiently cooperative or collaborative manner, each may, without prejudice to our other rights and remedies, temporarily suspend partners from involvement. IFOW could terminate participation where:

- partners committed a material or repeated breach of the Terms of Participation;
- IFOW determined partner conduct, either in the course of or outside of the Sandbox, is contrary to the public interest or is likely to bring IFOW or the Regulator Plus Forum into disrepute, or is otherwise deemed by us to be contrary to the nature of the Sandbox including where in our opinion the partner withhold information where they should reasonably have disclosed it, or where they took material action concerning the Proposed Innovation without informing or involving us.

# 4. Methodology

---

We present here both our approach to analysis of the data, and our approach to its collection in the first phase of operations.

We note that our Methodology within the Sandbox can and will change over time, according to the needs of research in this space.

## Approach to analysis

Our objective is to involve multiple cases in each Open Call. However, even single cases can provide significant insight into how law and governance are working, or could work, in practice. We abided by several key design principles, derived from frameworks developed to understand workplace change, which allow for broader lessons from fewer, or even single cases, using the extended case study method.<sup>32</sup>

The extended case method involves abiding by the following key principles in the treatment of data collected.

1. **Theoretically Oriented:** The case is selected because it relates to a theoretical puzzle or debate, not because it is “representative.” Its purpose is to refine, challenge, or extend theory.
2. **Situated in Wider Contexts:** The case is not treated as self-contained. It is embedded in larger social structures (economy, politics, culture) and historical processes.
3. **Attentive to Anomalies and Contradictions:** Instead of smoothing over inconsistencies, a good case highlights tensions that push theory forward. “Trouble” in the data becomes the basis for insight.
4. **Reflexive:** The researcher acknowledges their own role and impact on the field, and how the social situation or context (in this case, a Sandbox operated by an independent research and development institute) shapes knowledge production.
5. **Processual and Historical:** The case is traced over time, showing how social processes unfold and transform. Attention is given to temporality, not just a static snapshot.
6. **Extendable Beyond Itself:** The case provides leverage for thinking about broader dynamics. Its insights “travel” to explain patterns in other settings, even though it is not statistically generalisable. This comes through attention to how the case sits within wider historical and geographic context.

The application of these principles can be seen in our Phase 1 outputs, ‘Value and Values in AI Adoption’ and ‘Post-Deployment Training Data as the AI Existential Threat’.

## Approach to Data Collection

Good Work Algorithmic Impact Assessment<sup>33</sup> is aligned with international precedents for algorithmic impact assessment, tailored to the context of the workplace.<sup>34</sup> While principally designed as a framework for collaborative, responsible innovation that centres participation of affected stakeholders in a process that focuses on surfacing information and making decisions, with these recorded in a set of key documents.<sup>29</sup> Because it is designed to invite reflection on key choices throughout the design, development and deployment process, it surfaces information about both risks (ex-ante) and impacts (ex-post) of new tools. While the focus is on employers or the ‘enterprise’ level, the process looks ‘up’ to SaaS and PaaS relationships, allowing our ‘full stack’ approach.

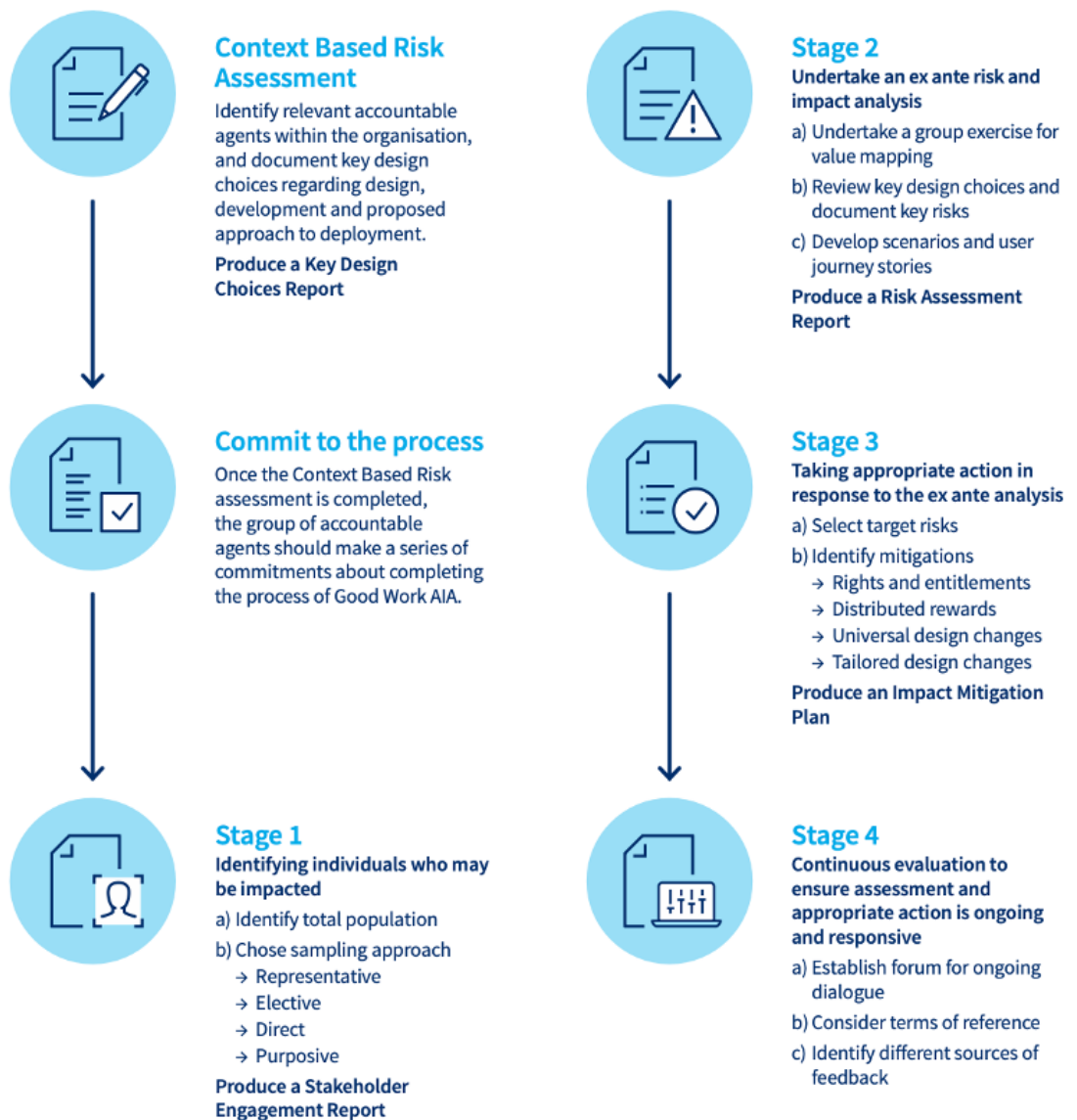
This allows us to surface:

- The **feasibility and utility of Good Work Algorithmic Impact Assessment**, and refinement of methods and conditions for success.
- The **risks and impacts of novel technologies**, across different design, development and deployment choices, sectors, firm sizes, corporate governance structures and their mitigation.
- The interoperability and interpretation of hard and soft regulatory regimes, as linked to interpretation, roles, professions, practices, and philosophies.

## Process for the GWAIA

### Set Up Phase - Documenting Key Decisions

Graphic 3 - Good Work Algorithmic Impact Assessment Stages



This preliminary stage requires firm to document choices about:

### Design

- A poor approach to problem definition and understanding relationships between technical and social systems is likely to obstruct productivity gains and increase negative over positive impacts on Good Work.
- Enterprises are encouraged to document the problem to be solved by technology and rationale for this.
- The intended or desired new forms of value creation from adoption, as these link to different automation archetypes (creation, displacement, high discretion augmentation, low discretion augmentation, intensification, matching, telepresence).<sup>35</sup>
- Explicit or implicit expectations about changes to job or task design, as these underpin different automation archetypes.

### Development

- Whether built in-house, or procured, the process suggests securing a summary description of the algorithmic system including its capabilities, the outcomes and optimisation functions, training methodologies and datasets used, the relative weighting of different variables, the techniques used to test and validate the system, trade-offs made between different measures, and recognition of policies relevant to the procurement of the system during processes of audit and rationale.
- If the system is developed and/or operated by a third-party, additional information should be ascertained, including the name of the third-party contracted, name of software, policies and processes relevant to procurement, and relationships and procedures of accountability.

### Deployment

When planning implementation, conscious choices should be made about the proposed integration of an algorithmic system within the social infrastructure of a business (team structure, job design, hierarchies of responsibility). Decisions about this should be made explicit: physical integration - for instance, how will the system be embedded in working life, and what hardware will it use?; Social integration - for instance, who will have access to or view recommendations or reports from the system; and oversight - for instance, who is the named accountable agent for the system?

**This produces a ‘Key Disclosures Report’.**

### The Impact Assessment Phase

Before the Impact Assessment starts, firms are required to make a series of commitments. Firstly, documenting their baseline commitments to good work as an organisation. Secondly, placing trust in a participatory process. Then making a commitment to resourcing responsible innovation (including time of staff, costs of mitigations, and learning capability). Firms are then required to make commitments to mitigations at different thresholds before the impact assessment begins.

#### Stage 1: Identifying impacted individuals

This involves different methods for identifying and then sampling, for later involvement, individuals have a freedom, right or interest which may be impacted by the system, including all workers (not only employees). Representative, Elective, Direct and Purposive samples are aligned with different later methodologies.

**This stage produces a Stakeholder Engagement Report.**

## Stage 2: Ex-Ante Risk Assessment

Drawing from information from Stage 1, different methods for ex-ante risk and impact assessment are undertaken with workers. This considers all ten good work dimensions across firms of different size, corporate governance structure, etc. This involves an initial value mapping, preliminary review of the Key Disclosures Report to produce 'constructive design commitments'; and scenario development.

**This produces a Risk Assessment Report.**

## Stage 3: Taking Appropriate Action

This involves determining levels of risk, and selecting priority risks to develop interventions for, or whether certain risks trigger moves to cease to implement. Then, identifying and developing mitigations (which could include new rights and entitlements, distributed rewards, universal reprogramming, or tailored reprogramming).

## Stage 4: Continuous Evaluation

Setting up a forum to oversee the implementation of the system and review effectiveness of mitigations.

Overall: how various hard and soft regulatory regimes shape outcomes/decision making (as shaped by interpretation and understanding, AI literacy, professional codes of conduct, power dynamics).

# 5. Disclosures: Funding

Each time we publish a new methodology report, we will also share developments in our funding. This is critical for upholding appropriate and fair scrutiny and abiding by our transparency principles.

As a civil society organisation, our work has been funded by philanthropic sources. We are grateful to the following organisations for their support in getting the survey to this point.

We also benefited from pro bono legal advice from Lewis Silkin on the development of our Terms of Participation, for which we are deeply grateful. We are also grateful to Jodi Starkman of the IRC4HR for her advice and guidance.

We ensure anonymity for all participating firms. This is to ensure we do not provide incidental private benefit to those firms in terms of brand recognition. However, we are very grateful for their time.

Funder	Activity	Period
<b>UK Information Commissioner's Office</b>	Support to develop the Good Work Algorithmic Impact Assessment	2020-2021
<b>Joseph Roundtree Charitable Trust</b>	Establishing the Regulator Plus Forum	2022-2024
<b>Friends Provident Foundation</b>	First Open Call, and research into the impacts of the Key Disclosures Stage and awareness of Automation Archetypes in decision making.	2024-2025
<b>Trust for London</b>	Sandbox set up, and improving the involvement and representation of low-income Londoners in the development of AI governance	2025-2027

## 6. Conclusions

---

The UK's longstanding leadership in AI has been paralleled by its role as a pioneer of the now-popular approach to its experimental governance: sandboxing. Now is an appropriate time to continue this global leadership by redefining how these architectures can operate to steer, shape and drive the transition to a new economy.

To remain competitive and relevant while harnessing the power of AI, the UK must take a broader approach than deregulation, just as it must take a more ambitious approach to new value creation through AI adoption than the pursuit of efficiencies. To steer the capabilities of new technology in the service of effective and sustainable growth, approaches which surface the interaction of technical and human capabilities – as these relate to and are shaped by compliance with the law, soft and informal rules, and professional norms and standards – are critical.

Sandboxing has the potential to achieve this, by using a framework and methodology which allows recognition of the interactions between these factors and a process which looks to promote innovation and good work together. Our research suggests that good work is a mediator and precondition for meaningful productivity.

The domain in which all AI tools are adopted and have the potential to afford meaningful productivity is the workplace. This is a multifaceted legal, regulatory and cultural domain. When sandbox environments are restricted to the application of a single regime, the ability to comprehend relationships between decisions, behaviour, interpretation, and interoperability is weakened. To effectively redesign our legal system to support innovation, we must understand how competing demands come into play in the everyday practices – and regionally varied contexts – of business decision-making. To understand which laws can be practically suspended or deleted, surfacing their intersections and interdependencies is also key.

# References

1. EU AI Act Article 57: AI Regulatory Sandboxes. Date of entry into force: 2 August 2026. According to: Article 113 Available at: <https://artificialintelligenceact.eu/article/57/> Accessed 15th April 2026.
2. US Sandbox Act, S.2750 — 119th Congress (2025-2026) Available at: <https://www.congress.gov/bill/119th-congress/senate-bill/2750/text> Accessed 15 April 2026.
3. Department for Science, Innovation and Technology. AI Growth Lab Call for Evidence. <https://www.gov.uk/government/calls-for-evidence/ai-growth-lab> Accessed 15 April 2026.
4. IFOW, The Good Work Charter. See toolkit summary of relevant regulation, at: <https://www.ifow.org/toolkit/the-good-work-charter> Accessed 15 April 2026
5. Soffia, Magdalena, Rolando Leiva-Granados, Xingzuo Zhou, and Jolene Skordis 'Does technology use impact UK workers' quality of life? A report on worker wellbeing, Institute for the Future of Work, February 2024.
6. Institute for the Future of Work (2025). Final Report of the Pissarides Review into the Future of Work and Wellbeing. London: Institute for the Future of Work. DOI: 10.5281/zenodo.14615712.
7. Daron Acemoglu. "Why nations fail?." The Pakistan Development Review 54, no. 4 (2015): 301-312.
8. Eldeman Trust Institute (2024) 'Edelman Trust Barometer 2024' Available at: <https://www.johnsmithcentre.com/news/research-uks-trust-in-government-among-the-lowest-globally/>
9. National Centre for Social Research (NCSR) (2024). Trust and confidence in Britain's system of government at record low. Available at: <https://natcen.ac.uk/news/trust-and-confidence-britains-system-government-record-low>. Accessed 16th April 2026.
10. IFOW (Forthcoming as of April 2026). 'Governing the Innovation Ecosystem: Differentiated Impacts and the Role of Rights'.
11. IFOW (Forthcoming as of April 2026) Automation Archetypes: Legal Protections and Gap Analysis,
12. Johnson, W. G. (2022). Caught in quicksand? Compliance and legitimacy challenges in using regulatory sandboxes to manage emerging technologies. Regulation & Governance, 17(3), 709–725.
13. The World Bank Group, (2020) 'Global Experiences from Regulatory Sandboxes' <https://documents1.worldbank.org/curated/en/912001605241080935/pdf/Global-Experiences-from-Regulatory-Sandboxes.pdf> Accessed 15th April 2026.
14. Goo, Jayoung James, and Joo-Yeun Heo. "The impact of the regulatory sandbox on the fintech industry, with a discussion on the relation between regulatory sandboxes and open innovation." Journal of Open Innovation: Technology, Market, and Complexity 6, no. 2 (2020): 43.
15. Renieris, E. M. et al. (2023). Building Robust RAI Programs as Third-Party AI Tools Proliferate Published by Boston Consulting Group (BCG) and MIT Sloan Management Review. Available at: [https://web-assets.bcg.com/1b/18/c684f0174e088e068efc4c62c942/building-robust-rai-programs-as-third-party-ai-tools-proliferate.pdf?utm\\_source=chatgpt.com](https://web-assets.bcg.com/1b/18/c684f0174e088e068efc4c62c942/building-robust-rai-programs-as-third-party-ai-tools-proliferate.pdf?utm_source=chatgpt.com) Accessed 15th April 2026.
16. Gilbert et al., (2026) 'Economic Sovereignty and the Question of Post Deployment Training Data' Institute for the Future of Work. Available at: [www.ifow.org/publications/economic-sovereignty-and-the-question-of-post-deployment-training---a-research-report-from-the-ifow-sandbox](http://www.ifow.org/publications/economic-sovereignty-and-the-question-of-post-deployment-training---a-research-report-from-the-ifow-sandbox).
17. Pekdemir, Ceren. "On the regulatory potential of regional organic standards: Towards harmonization, equivalence, and trade?." Global Environmental Change 50 (2018): 289-302.
18. IFOW Regulator Plus Forum. Available at: <https://www.ifow.org/about-us/our-networks>. Accessed 16th April 2026.
19. Hayton, J. et al. (2023) What drives UK firms to adopt AI and robotics, and what are the consequences for jobs? Institute for the Future of Work. Available at: <https://www.ifow.org/publications/adoption-of-ai-in-uk-firms-full-working-paper> Accessed 15th April 2026
20. Cooper, Davina. "Towards an adventurous institutional politics: The prefigurative 'as if' and the reposing of what's real." The Sociological Review 68, no. 5 (2020): 893-916.
21. Yates, L. (2015). Rethinking prefiguration: Alternatives, micropolitics and goals in social movements. Social Movement Studies, 14(1), 1–21.
22. All Party Parliamentary Group on the Future of Work (2021) 'The New Frontier: Artificial Intelligence at Work' [https://cdn.prod.website-files.com/64d5f73a7fc5e8a240310c4d/64d5f73b7fc5e8a240311348\\_The%20New%20Frontier%20-%20Artificial%20Intelligence%20at%20Work%5B4%5D.pdf](https://cdn.prod.website-files.com/64d5f73a7fc5e8a240310c4d/64d5f73b7fc5e8a240311348_The%20New%20Frontier%20-%20Artificial%20Intelligence%20at%20Work%5B4%5D.pdf) Accessed 15 April 2026.
23. Kay, Jackie, Atoosa Kasirzadeh, and Shakir Mohamed. "Epistemic injustice in generative AI." In Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, vol. 7, no. 1, pp. 684-697. 2024.
24. Burawoy, Michael. "Revisits: An outline of a theory of reflexive ethnography." American sociological review 68, no. 5 (2003): 645-679.
25. Jakesch, Maurice, Zana Bućinca, Saleema Amershi, and Alexandra Olteanu (2022). "How Different Groups Prioritize Ethical Values in Responsible AI." arXiv preprint arXiv:2205.07722.
26. A survey of workers conducted by IFOW in 2021 found 52 percent of workers were "not at all confident" that they knew "why and for what purposes" their employers used data collected about them, and 67 percent were "not at all confident" that they understood how their data was used to assess their performance. See findings in: Gilbert, Abigail and Anna Thomas, "The Amazonian Era: The Gigification of Work." Institute for the Future of Work. (2021) Available at: <https://www.ifow.org/publications/the-amazonian-era-the-gigification-of-work>.

27. Fricker, Miranda. "Epistemic justice as a condition of political freedom?" *Synthese* 190, no. 7 (2013): 1317-1332.
28. Ulbricht, Lena, and Karen Yeung. "Algorithmic regulation: A maturing concept for investigating regulation of and through algorithms." *Regulation & Governance* 16, no. 1 (2022): 3-22.
29. Vardi, Moshe Y. "Move fast and break things." *Communications of the ACM* 61, no. 9 (2018): 7-7.
30. Chen, Gabriel Kwok Hui, and Araz Taeihagh. "Designing regulatory sandboxes: a comprehensive framework for aligning functionalities and objectives." *Policy Design and Practice* 9, no. 1 (2026): 1-15.
31. Ranchordas, Sofia, and Valeria Vinci. "Regulatory sandboxes and innovation-friendly regulation: Between collaboration and capture." *Italian J. Pub. L.* 16 (2024): 107.
32. Burawoy, Michael. "The extended case method." *Sociological theory* 16, no. 1 (1998): 4-33.
33. Gilbert, Abigail and Anna Thomas. 'Good Work Algorithmic Impact Assessment'. Available at: [https://www.ifow.org/publications/good-work-algorithmic-impact-assessment-an-approach-for-worker-involvement?gad\\_source=1&gad\\_campaignid=22641629958&gbraid=0AAAAA-OijA0fyWW-qLq3Wq55VBcbhxkZ&gclid=CjwKCAjw-dfOBhAjEiwAq0Rw1uTM6Ew8ap-jYJs-GxR7faCMxcOAFdaSMYxNR37xkkQN0DtAV3DeRoCBbAQAvD\\_BwE](https://www.ifow.org/publications/good-work-algorithmic-impact-assessment-an-approach-for-worker-involvement?gad_source=1&gad_campaignid=22641629958&gbraid=0AAAAA-OijA0fyWW-qLq3Wq55VBcbhxkZ&gclid=CjwKCAjw-dfOBhAjEiwAq0Rw1uTM6Ew8ap-jYJs-GxR7faCMxcOAFdaSMYxNR37xkkQN0DtAV3DeRoCBbAQAvD_BwE).
34. Council of Europe Committee on Artificial Intelligence (2024) 'Methodology for the Risk and Impact Assessment of Artificial Intelligence Systems from the Point of View of Human Rights, Democracy and the Rule of Law' Available at: <https://rm.coe.int/cai-2024-16rev2-methodology-for-the-risk-and-impact-assessment-of-arti/1680b2a09f>
35. Gilbert, Abigail (2024) 'Reframing Automation: A New Model for Anticipating Risks and Impacts' Institute for the Future of Work, Available at: <https://www.ifow.org/publications/reframing-automation-a-new-model-for-anticipating-risks-and-impacts> Accessed 15th April 2026.



# Institute for the Future of Work

IFOW is an independent research and development institute dedicated to transforming working lives for the better, co-founded by former employment barrister Anna Thomas MBE, Nobel prize-winning economist Sir Christopher Pissarides, and technologist Naomi Climer CBE.

Our core team at Somerset House works with a growing network of strategic partners striving for systems change.

**Our vision** is a future in which everyone flourishes in work they shape.

**Our mission** is to understand together how to transform working lives for good.

**Our theory of change** is that creating and sustaining good work is the best way to achieve this goal and ensure that innovation and social good advance together.



Trust for London is one of London's largest funders, and has been supporting work for a fairer city since 1891. Their stated belief is that 'all people have the right to reach their full potential' and they fund organisations working at the frontline of social and economic justice.

Institute for the Future of Work  
Somerset House  
Strand  
London  
WC2R 1LA  
[ifow.org](http://ifow.org) / [@ifow.org](https://twitter.com/ifow.org)

## About our Reports

Our expert Reports present new work from across IFOW's research team, research fellows, practitioners and partners, offering policymakers, firms and academics access to the latest research in the Future of Work space.

If you would like to further information, please contact [team@ifow.org](mailto:team@ifow.org)

## Citation

Gilbert, A., Marriott, J., Leslie, M., *IFOW Sandbox - Phase 1 Methodology Report*. London: Institute for the Future of Work. DOI: 10.5281/zenodo.18480973

## Permission to share

This document is published under the Creative Commons Attribution Non Commercial No Derivatives 4.0 England and Wales Licence. This allows anyone to download, reuse, reprint, distribute, and/or copy IFOW publications without written permission subject to the conditions set out in the Creative Commons Licence.