AI for Fair Work

From principles to practices

December 2023 – GPAI New Delhi Summit

GPAI / THE GLOBAL PARTNERSHIP ON ARTIFICIAL INTELLIGENCE

Table of Contents

Executive Summary

The GPAI Future of Work group has translated the OECD AI principles into concrete workplace standards to inform the practices of employers. This process of standard development initially began with a global stakeholder consultation, which generated a first version. Now, the research team have conducted two in-depth workplace studies to connect these standards to empirical data collected on the frontlines of AI system deployment. This process has allowed for a revision of the principles, conducted in association with research partner Fairwork. The new principles have a stronger focus on pay, conditions, contracts, management and representation, because these fundamental issues were consistently highlighted by research participants. A second report containing the results of the two case studies is to follow in early 2024.

Introduction

The AI policy environment is developing rapidly. Debates around the appropriate deployment of AI systems have progressed significantly from the publication of our last report, as statements of principle continue to be progressively replaced by more binding regulation. Governments have adopted various models of AI regulation, leading to divergent paths of potential technological development. As new ILO research shows, the dominant regulatory approaches can be divided into three schools: technological leniency, hard ex-ante regulation and soft-law.[1] The contradictions between these approaches has generated significant opacity for consumers, developers and regulators. The GPAI Future of Work group's approach has never been to advocate for any one of these schools, but rather to develop specific concrete workplace standards that translate the OECD's principles into the realm of practice. The goal of this report is to update readers on the current process of principle development, and highlight the future work that the project team will undertake to amplify the impact of these principles in practice. Crucially, this report will be complimented by a significant subsidiary report in 2024 that will present the results of the case studies conducted by the project and show how these principles have been applied to significant effect in practice.

Previous work

In 2019, the OECD adopted the Recommendation of the Council on Artificial Intelligence. This was a major landmark in the international development in the discussion and regulation of AI ethics. The recommendation covered five complimentary valued-based principles:

1. Inclusive growth, sustainable development and well-being
2. Human-centred values and fairness
3. Transparency and explainability
4. Robustness, security and safety
5. Accountability

These principles had significant implications for the future of work. Workplace deployments of AI in, (for example, digital labour platforms) were already significant and were only becoming more so over time. It was in this context that the GPAI Future of Work Working Group decided to launch a new project titled AI for Fair Work in October 2021. The project took the 2019 Recommendations as a starting point, and aimed to develop them into concrete standards applicable to the workplace through a global tripartite consultation with governments, regulators, international bodies, corporations, trade unions, and experts. This consultation consisted of two rounds of stakeholder engagement: first, the

---

[1] Ekkehard Ernst et al., "The Fight for AI Regulation: A Global Political Economy Perspective," in *The Fight for AI Regulation: A Global Political Economy Perspective* (IPSA WC 2023, Buenos Aires, 2023), https://wc2023.ipsa.org/wc/paper/fight-ai-regulation-global-political-economy-perspective.

team conducted 21 interview and one focus group with key global AI leaders; second, the team circulated a survey to 117 further stakeholders to elicit their feedback on the final draft of the principles. The results of this consultation were published in 2022.[2]

1. Guarantee fair work: Ongoing changes in work caused by the introduction of AI systems have the potential to disrupt the labour market, but internationally agreed minimum rights and standards remain a precondition of fair AI.

2. Build fair production networks: AI system development and deployment relies on global networks of human labour, hardware production, and infrastructure. Organisations seeking to implement fair AI in the workplace must therefore look beyond the immediate production process to the networks of production that enabled it and use their procurement power to achieve fairness across the network.

3. Promote explainability: Workers have a right to understand how the use of AI impacts their work and working conditions. Organisations must respect this right and provide detailed, understandable resources to allow workers to exercise it.

4. Strive for equity: AI systems have been found to reproduce and scale up patterns of social discrimination. The costs associated with embedding negative consequences for marginalised groups into workplace technology are extremely high. As a result, AI systems must be (re)designed, built, and deployed in a way that actively seeks to eliminate sources of discrimination. Processes such as audits and impact assessments should be integrated into the AI system lifecycle to allow for ongoing scrutiny.

5. Make fair decisions: the automation of decision making can lead to reductions in accountability and fairness. But building in human oversight into a decision making loop doesn't solve this problem. Instead, the subjects of those decisions need to be empowered to challenge them, and a renewed emphasis should be placed on the liability of those stakeholders who direct the development and deployment of AI systems in the workplace.

6. Use data fairly: the collection of large quantities of data and the concentration of its ownership may exacerbate risks for individuals and social groups, especially when shared with third parties. Limits must therefore be put on collection (i.e. data minimisation) and processes must be instituted for subjects to access and protect their data in a comprehensive and explainable format. Organisations should provide comprehensive guidelines for individuals to understand data ownership, data usage and any potential risks that result, so that they are able to question, contest, and when necessary, reject, decisions made about them.

7. Enhance safety: advances in algorithmic management have increased the risks of work intensification and surveillance. In this context, the right to healthy, safe working environments must be protected. Potential improvements in safety should be capitalised upon, but deployment must take place in a way which reflects the different understandings of stakeholder groups about the trade-offs involved.

8. Create future-proof jobs: the introduction of AI systems to workplaces can cause specific risks such as job destruction and deskilling. These risks can be reduced by treating the introduction of AI as an opportunity for workers and organisations to engage in a participatory and evolutionary redesign of work which uses the rewards of AI to increase job quality.

9. Avoid inappropriate deployment: organisations should proactively test AI systems to a high standard in order to avoid harms in advance, rather than iterating to address them post-deployment.

10. Advance collective worker voice: the risks and rewards of AI systems are understood differently by different stakeholder groups. These divergences should be proactively

---

[2] Callum Cant et al., "AI for Fair Work" (Paris: Global Partnership on Artificial Intelligence, 2022), https://www.gpai.ai/projects/future-of-work/AI-for-fair-work-report.pdf.

negotiated, rather than suppressed. Pursuing AI system implementation in a multi-stakeholder environment requires a mechanism to turn ethical principles into ethical practice through democratic participation by workers. Collective bargaining between workers and management is best suited to play this role.

Project overview

The 2022-23 cycle of the AI For Fair Work project has partnered with Fairwork, a research project based at the University of Oxford and the Berlin Social Science Centre, to further develop the GPAI principles into a highly-concrete framework. The research team used real world data to ensure that the principles respond effectively to a range of geographical, industrial and social contexts. To do so, we have conducted two qualitative case studies: one focused on Amazon (the massive online retailer) in the UK, the other on Sama (a data annotation firm specialising in computer vision) in Kenya and Uganda. We also collected supplementary data from Japan.[3]

Both case studies were based on extensive fieldwork. The exact method varied slightly due to the degree of access granted by the subject of the case study to the research team, but the research methods deployed in both studies involved participant observation of work and worker organising, alongside semi-structured interviews with employees and management. In both cases, we collected a significant quantity of data on specific workplace uses of AI systems and conducted a thematic analysis to establish both how the principles could be revised to reflect the specific questions of fairness that arose in the determinate context of the case studies, and how far the workplaces studied met the basic thresholds of fairness we established. This initial data analysis has been followed, in the case of Sama, by extensive cooperative engagement from corporate leadership to make progress towards meeting the GPAI principles in practice. We believe that these case studies will provide significant reference points for the implementation of fairer AI in the workplace which can be cited by a wide range of actors seeking to form policy or best practice. In addition, they act as a pilot for further work by the project in which we will gather data on a wider range of companies and implement a modified version of the Fairwork action research method to actively advocate for the implementation of the GPAI standards in industry. The full results of the two case studies are not yet available for publication due to ongoing consultation with the two subjects of the case studies. They will be published by Fairwork in 2024.

What next

Now, building upon insights from our 2023 East African fieldwork, the project aims to evaluate working conditions in the global AI data pipeline, and identify pathways for positive change by engaging with policymakers, lead AI firms, client BPO firms, and workers' associations to promote fair work practices aligning with the GPAI's established guidelines.

The next phase of the project will see us conduct a wider investigation into the working conditions of data annotation workers in the AI data pipeline with a specific focus on Latin America, in particular Mexico, Argentina and Colombia. These countries all act as global hubs for the data annotation industry, which sells services to client firms mostly based in the Global North.

The study will employ a qualitative methodology, using 45 worker interviews, supplemented by desk research and management interviews, to examine the working conditions of data annotators at 9 firms. Workers and company representatives will be interviewed to gain insights into their experiences and practices. The 9 data annotation firms studied will be awarded a score out of 10 based on the evidence gathered. This method design builds on 5 years of experience studying digital labour platforms in 38 countries as part of the wider Fairwork project.

---

[3] The research team would like to specifically thank GPAI expert Kyoko Yoshinaga for her help with this task.

The next phase of the project will have three key research objectives:

1. Assess Working Conditions in the AI Data Pipeline: Examine the working conditions of data annotation workers within the AI data pipeline, particularly in the Global South. This includes identifying both fair and unfair labour practices and developing scores out of ten for 9 BPO firms using the Fairwork AI principles as a benchmark. This assessment should be communicated to the broadest possible audience.

2. Understand Connections to Lead Firms: Investigate the relationships between these workers and leading tech firms in the Global North. This includes tracing the flows of resources, data, and capital that link them.

3. Promote Fairness in the AI Data Pipeline: Advocate for fairer conditions for data annotators in the BPO industry. By scoring firms out of 10 we will incentivise positive changes. The linked lead firms that contract data annotation to the scored BPO firms will be positively incentivised to support pro-fairness changes in their supply chains.

Revised Principles

The major change between the principles between version one (produced during the 21-22 cycle) and version two (produced during the 22-23 cycle) is an much greater focus on the basic issues of pay, conditions, contracts, management, and representation. During fieldwork, we found that that the issues highlighted in version one did not reflect the primary concerns of workers on the ground. Whilst they may have reflected the priorities of global, policy-orientated stakeholders, they did not always overlap with the fundamental concerns raised in interviews and during observations. This insight led the research team to engage in a reflective process, through which we decided that the over-specialisation of version one could be corrected by rewriting the principles in line with other existing global standards. As a collaborator on the project, the Fairwork team highlighted the effectiveness of their five principle model, and proposed adapting the standards embedded in version one into a modified form of the Fairwork principles to create a version two that was more responsive to the concerns that emerged through the process of data collection. The result is the following set of AI principles:

FAIRWORK AI Principles

The following set of principles are intended to be used to assess the use of AI systems in the workplace in the context of the employment relation. They build on the AI for Fair Work principles established by the GPAI in 2022 on the basis of a global multistakeholder consultation.

1. Fair Pay

1.1 Pays at least the local minimum wage

To achieve this point, the employer takes appropriate steps to ensure ALL of the following:

- Workers, regardless of their employment status or contract type, must earn the local minimum wage[4] or the wage set by collective sectoral agreement (whichever is higher) for all hours worked.[5]

- Workers, regardless of their employment status or contract type, are paid on time and in-full.

1.2 Pays at least the local living wage

Minimum wage can be insufficient to ensure workers and their dependents a basic but decent standard of living. The living wage exists to set the benchmark of what is required to enable this decent standard of living.[6]

To achieve this point, the employer takes appropriate steps to ensure the following:

- Workers, regardless of their employment status or contract type, must earn at least the living wage, or the wage set by collective sectoral agreement (whichever is higher) for all hours worked.

---

[4] The ILO defines minimum wage as the "minimum amount of remuneration that an employer is required to pay wage earners for the work performed during a given period, which cannot be reduced by collective agreement or an individual contract." Minimum wage laws protect workers from unduly low pay and help them attain a minimum standard of living. The ILO's Minimum Wage Fixing Convention, 1970 C135 sets the conditions and requirements of establishing minimum wages and calls upon all ratifying countries to act in accordance. Minimum wage laws exist in more than 90 per cent of the ILO member states.

[5] This means not only that the rate of pay agreed with workers reaches that statutory level, but also that workers are accurately compensated for all hours worked. Underpayment (also known as 'wage theft') is a pervasive problem, with evidence suggesting that huge sums of value go unpaid due to unpaid overtime, and incomplete/inaccurate wage payments.

[6] Where a living wage does not exist, Fairwork will use the Global Living Wage Coalition's Anker Methodology to estimate one.

2. Fair Conditions

2.1 Ensures safe working conditions

Workers face several risks in the course of their work, including strain, exhaustion, and exposure to traumatic content. They have a right to protection from these risks.[7] Employers must show they are aware of task specific risks and take steps to mitigate them.

To achieve this point, the employer must satisfy ALL of the following:

- Implement policies and practices that protect workers' safety from task specific risks. This should, at a minimum, account for well-evidenced risks such as:

- High job strain, which can lead to a range of negative health impacts including cardiovascular disease and mental health disorders.

- Secondary traumatic stress, which can be associated with repeated exposure to traumatic content.

- Muscular skeletal injuries, which may emerge as a result of unsuitable equipment, excessive workload or perverse incentivisation in physical jobs.

- Risks related to a specific job are flagged to workers before they accept the job (such as indicating that they might be exposed to violent content.)

- The employer places a maximum limit on standard working time that meets either the applicable national regulation or, in cases where there is no applicable national regulation, the ILO standard of 40 hours a week.[8]

- Workers are entitled to take breaks during working time that is defined under the applicable national regulation, or in cases where there is no applicable national regulation, is equivalent to a minimum of one hour for every eight hours worked.

- If the work arrangements require workers to work in shifts, workers are given the option to choose their shifts, and reasonable accommodations are made for workers with additional needs due to health, safety and other personal reasons (such as pregnancy, care requirements, disability and other health conditions.)

2.2 Ensures paid leave, and a safety net

Workers are vulnerable to the possibility of losing their income as the result of unexpected or external circumstances, such as sickness or injury. Most countries provide a social safety net to ensure workers don't experience sudden poverty due to circumstances outside their control. However, not all workers might qualify for the social safety protections due to their own personal circumstances (e.g. visa status, residency status). In recognition of the fact that most workers are dependent on income they earn from the work, employers must ensure that workers are compensated for loss of income due to inability to work. In addition, employers must minimise the risk of sickness and injury.

To achieve this point, the employer must ensure ALL of the following:

- Workers have access to paid time-off (such as bereavement, parental, sick and annual leave.)

---

[7] The ILO recognises health and safety at work as a fundamental right. Where the platform directly engages the worker, the starting point is the ILO's Occupational Safety and Health Convention, 1981 (C155). This stipulates that employers shall be required "so far as is reasonably practicable, the workplaces, machinery, equipment and processes under their control are safe and without risk to health", and that "where necessary, adequate protective clothing and protective equipment [should be provided] to prevent, so far as is reasonably practicable, risk of accidents or of adverse effects on health."

[8] As endorsed by the ILO's Forty-Hour Week Convention, 1935 (No.47) and the Reduction of Hours of Work Recommendation, 1962 (No.116)

Where core medical treatment is not provided by a public system, such as a national healthcare scheme, the employer makes a meaningful provision to the health care costs of its workers.

3. Fair Contracts

3.1 Provides decent contracts

Employment on temporary contracts can have significant negative effects on job satisfaction, well-being and health. Short-term contracts, such as those lasting one to three months or with no guaranteed working hours, place workers in precarious positions and are likely to exacerbate these negative effects.

To achieve this point, the employer must meet ALL of the following:

- Workers must sign a contract and/or give informed consent to terms of conditions upon signing up, and for each subsequent contract extension.

- The contract or terms and conditions is presented in full, in clear and comprehensible language that all workers could be expected to understand.

- The contract or terms and conditions are easily accessible to workers in paper and/or electronic form. If these conditions differ for different contract types, reasonable steps are taken to inform workers about the differences in contract types.

- The party employing the worker must be identified in the contract or terms and conditions, and subject to the law of the place in which the worker works.

- Workers working on long-term projects that exceed the probation time are provided with the option to sign an employment contract lasting at a minimum the same length of time as the project.

- The contracts or terms and conditions do not include clauses that revert prevailing legal frameworks in the countries where workers work.

3.2 Provides secure employment

Whilst fixed-term employment may be suitable for some workers' circumstances, secure employment is a fundamental improvement of working conditions for many others.

To achieve this point, the employer must meet ALL of the following:

- Workers with three years or more of consistent short-term employment should be provided with the option to move onto permanent contracts if they so desire.

- The employer should make reasonable adjustments in wages and conditions between both: fixed-term and permanent employees and outsourced workers; and any outsourced or indirectly employed workers and directly employed workers. Workers who are outsourced or indirectly employed should be compensated for additional costs incurred, including visa/work permits and their extensions, insurance, pensions, and other social security premiums.

- In cases of justified redundancy or contract non-renewal, the employer should provide workers with severance allowance commensurate with tenure at the company and retraining opportunities. In cases where the redundancies are being made because reasons of an economic, technological, structural or similar nature, workers or their representatives are consulted, and steps are taken to minimise the resulting redundancies.[9] If desired, workers should be able to invite worker representatives to their end of contract meetings with the relevant HR departments.

---

[9] The ILO Termination of Employment Convention, 1982 (No. 158) defines worker representative consultation as sufficient when the employer provides "the workers' representatives concerned in good time with relevant

In the case of subcontracting arrangements, where part or all of the work is subcontracted to other companies, management implements a reliable mechanism to monitor and ensure that the subcontractor is living up to the standards expected from the company itself regarding working conditions.

4. Fair Management

4.1 Treats workers fairly

The employment relation is an unequal one, with managers being afforded significant legal and economic sources of power not available to most workers. The interests of these two groups may diverge, leading to sometimes opposed immediate interests in the workplace. This dynamic can lead to unfair management practices.

To achieve this point, the employer must meet ALL of the following:

- Management should refrain from deploying any form of depersonalised bullying or mobbing in order to ensure organisational goals are met.[10]

- There is a policy in place which guarantees that any form of harassment in the workplace will not be tolerated.

- There is a policy in place which guarantees that the employer will not discriminate against persons on the grounds of racial, ethnic, social or minority background, caste, religion or belief, political or any other opinion, language, gender, gender identity, sex, sexual orientation, disability, age, geographical location, or any other status.

- Workers should have the right to appeal dismissals and other disciplinary measures.

- Workers are not disadvantaged for voicing concerns or appealing disciplinary actions.

4.2 Creates clear and effective systems for data management, explanations, and appeals

Contemporary workplaces are increasingly defined by data. The use of AI systems and automated management processes exacerbates both the incentives for employers to gather data from the work process, and diminishes the importance of workers' existing rights to receive explanations, appeal decisions, and access/own their data.

To achieve this point, the employer must meet ALL of the following:

- Where AI systems are involved in work, employers must create explainability mechanisms such as transparency reports or question and answer processes that allow workers to understand both the model behaviour of the system as a whole and specific decisions.[11]

- Workers must be able to appeal decisions made by AI systems through a multi-stakeholder process that reflects collective worker voice, and successful appeals to

---

information including the reasons for the terminations contemplated, the number and categories of workers likely to be affected and the period over which the terminations are intended to be carried out" and gives "in accordance with national law and practice, the workers' representatives concerned, as early as possible, an opportunity for consultation on measures to be taken to avert or to minimise the terminations and measures to mitigate the adverse effects of any terminations on the workers concerned such as finding alternative employment."

[10] Depersonalized bullying is a form of workplace mistreatment where employees are unfairly treated not because of who they are, but because of the organization's system or structure, constitutes a situation where harmful behaviour, like intimidation or aggression, are applied impersonally across the workforce by supervisors or managers in the name of achieving company goals.

[11] Workers have a right to understand how the use of AI impacts their work and working conditions. Organisations must respect this right and provide detailed, understandable resources to allow workers to exercise it.

lead not only that specific decision being revised but also wider revisions of decision-making process.[12]

- Management avoids excessive surveillance in the workplace, and avoids use of invasive technologies.

- Workers must not be subject to excessive data collection practices and should be informed about the data that is being collected about them. Employers must apply the principle of data minimisation (collecting the minimum amount of personal data required to fulfil a legitimate purpose) in their collection processes.

5. Fair Representation

5.1 Assures freedom of association and the expression of worker voice

Freedom of association is a fundamental right for all workers, and enshrined in the constitution of the International Labour Organisation, and the Universal Declaration of Human Rights. The right for workers to organise, collectively express their wishes – and importantly – be listened to, is an important prerequisite for fair working conditions.

To achieve this point, the employer must satisfy ALL of the following:

- There is a documented mechanism for the expression of collective worker voice that allows ALL workers, regardless of contract type or duration to participate in collective groups without risks.[13]

- There is a formal, written statement of willingness to recognise, and bargain with, a collective, independent body of workers or trade union, that is clearly communicated to all workers, and available on the company webpage. [11]

- Freedom of association is not inhibited, and workers are not disadvantaged in any way for communicating their concerns, wishes and demands to the company management, or expressing willingness to form independent collective bodies of representation.

5.2 Supports democratic governance

To realise fair representation, workers must have a say in the conditions of their work. This could be through a democratically governed cooperative model, a formally recognised union, or the ability to undertake collective bargaining with the employer.

To achieve this point, the employers must satisfy at least ONE of the following:

- Workers play a meaningful role in governing the company.

- In a written document available, the company publicly and formally recognises an independent collective body of workers, an elected works council, or trade union, and takes meaningful steps towards signing a collective bargaining agreement. This recognition is not exclusive and, when the legal framework allows, the company should recognise any significant collective body seeking representation.[14]

---

[12] The automation of decision making can lead to reductions in accountability and fairness. But building in human oversight into a decision-making loop does not solve this problem. Instead, the subjects of those decisions need to be empowered to challenge them, and a renewed emphasis should be placed on the liability of those stakeholders who direct the development and deployment of AI systems in the workplace.

[13] A mechanism for the expression of collective worker voice will allow workers to participate in the setting of agendas so as to be able to table issues that most concern them. This mechanism can be in physical or virtual form (e.g. online meetings) and should involve meaningful interaction (e.g. not surveys). It should also allow for ALL workers to participate in regular meetings with the management.

[14] If workers choose to seek representation from an independent collective body of workers or union that is not readily recognized by the platform, the platform should then be open to adopt multiple channels of representation, when the legal framework allows, or seek ways to implement workers' queries to its communication with the existing representative body

Conclusion

The GPAI Future of Work group's development of concrete principles of fair work has now gone through two cycles of iteration: an international stakeholder consultation, followed by an in-depth case study engagement. These processes have produced a set of principles that attempt to capture two perspectives: both those of high-level stakeholders, and the workers on the front lines of AI system deployment. This report has presented how the priorities of that second group prompted a significant revision of the principles in line with a general framework established by Fairwork. The resulting principles will now be deployed in further research, with a particular focus on the AI data pipeline in Latin America. Our goal, as a research team, is to use this next phase of research in 2023-24 to produce tangible results via stakeholder engagement using a tried-and-tested action research methodology.

The deployment of AI systems to the workplace is an ongoing, dynamic process. As researchers, engaging with that process demands reflexivity and an ability to dynamically redesign our interpretative frames as we gather more data. This latest phase of the AI for Fair Work project has proven that. Our future research will aim to maintain this reflexivity, and in particular to highlight the need for global stakeholders to foreground the voices and ideas of people on the frontlines of AI system deployment, whose experiences and concerns often diverge from the abstract priorities developed in high-level debates.