ETUI Policy Brief

European Economic, Employment and Social Policy

2021.07

The AI Regulation: entering an AI regulatory winter?

Why an ad hoc directive on AI in employment is required

Aída Ponce Del Castillo

Aída Ponce Del Castillo is a senior researcher at the Foresight Unit of the European Trade Union Institute (ETUI) in Brussels.

Policy issues

- AI systems in the context of employment are intrusive and have negative impacts on workers. The proposed Regulation fails to address the specificity of AI uses in employment, including platform work. An ad hoc directive on AI in employment is therefore necessary.
- As a consequence of focusing the regulatory approach on high-risk applications, the majority of use cases are considered low-risk, not subject to any evaluation and *de facto* authorised.
- High-risk uses are permitted subject to compliance with specific requirements and an *ex ante* conformity assessment based on internal control checks. This approach stacks the deck in favour of tech providers when the priority of this Regulation should have been to protect EU citizens and workers' rights.

etui.

Background

An AI regulatory package has been put forward by the European Commission's Executive Vice-President, Margrethe Vestager. This 'landmark' legislative initiative was announced by EC President, Ursula von der Leyen, in her 2019-2024 political guidelines. It follows the White Paper on AI and is presented as a fundamental building block of Europe's digital transformation and sovereignty.

In the AI draft Regulation, released on 21 April, the EC states that a legal framework on AI is needed to foster the development and uptake of AI and 'improve the functioning of the EU internal market, while meeting a high level of protection of public interests' (European Commission 2021). The starting point is clearly the desire to promote AI and this colours the whole regulation. The EC claims to be putting fundamental rights at the top of its approach but this concern systematically gives way to market-based priorities and the development of an industry that the EU intends to dominate and in which citizens and workers' rights are secondary. AI in employment should be taken out of the scope of this Regulation and made the subject of an ad hoc directive.

The building blocks of the EC's AI Regulation

A high-risk approach

The Regulation does not regulate AI as a technology but focuses on AI systems being placed on the market or put into service. It proposes a layered risk-based approach in which uses of AI that create unacceptable risk are prohibited; uses that create high risk are allowed where specific requirements are met; and uses with minimal or low risk are allowed, mostly unconditionally.

AI uses which are harmful to fundamental values are considered as unacceptably risky and prohibited. These are systems that deploy subliminal techniques, exploit vulnerabilities and distort human behaviour or are used for algorithmic social scoring. Using AI systems for 'real time' remote biometric identification of people in public spaces is considered particularly intrusive and prohibited except in three situations: the search for victims of crime; threats to life and terrorism; and identifying the whereabouts of perpetrators.

High-risk AI systems are at the heart of the Regulation. They are allowed on the market but must comply with certain mandatory requirements. Such systems may either be used as safety components within products or as one of eight types of stand-alone product: biometric identification and categorisation of people; management and operation of critical infrastructure; education and vocational training; essential private and public services (eligibility for benefits, creditworthiness, ...); law enforcement; migration, asylum and border control management; the administration of justice; and democratic processes. Importantly the field of employment is also considered a high-risk use.

Low-risk AI uses are not specifically addressed by the Regulation. Providers are simply encouraged to produce behavioural codes of conduct to foster the voluntary application of legal requirements appropriate to high-risk systems. Some other uses that are low risk such as AI systems that interact with humans, detect emotions or used to generate or manipulate audio, images or video content ('deep fakes') must meet specific transparency obligations: their provider should inform users that they are interacting with an AI system or that the content they are viewing has been manipulated.

What are the mandatory requirements for high-risk AI systems?

The Regulation establishes a system in which the production, sale and use of high-risk AI systems are permitted, within the existing scope of national and EU law, as long as they comply with a set of specific requirements and an *ex ante* conformity assessment. This assessment is based on internal control checks, so effectively what is being required is self-assessment. The only exception is remote biometric identification systems which are subject to a conformity assessment carried out by a third party.

The first of these specific requirements concerns datasets. High-risk AI systems which rely on the data-based training of models need to use high-quality datasets where the training, validation and testing is relevant, representative, free of error and complete. The second requirement is to establish a risk management system and to maintain this throughout the life cycle of the AI system. This should identify and analyse known, foreseeable risks while testing should be carried out to ensure systems work consistently with their intended purpose. The other requirements relate to data governance; technical documentation; record keeping (including the automatic generation of logs); transparency and the provision of information to users; human oversight; consistency of performance; robustness; accuracy; and cybersecurity and resilience. Before placing a high-risk AI system on the market, providers should register it in an EU database which is publicly accessible.

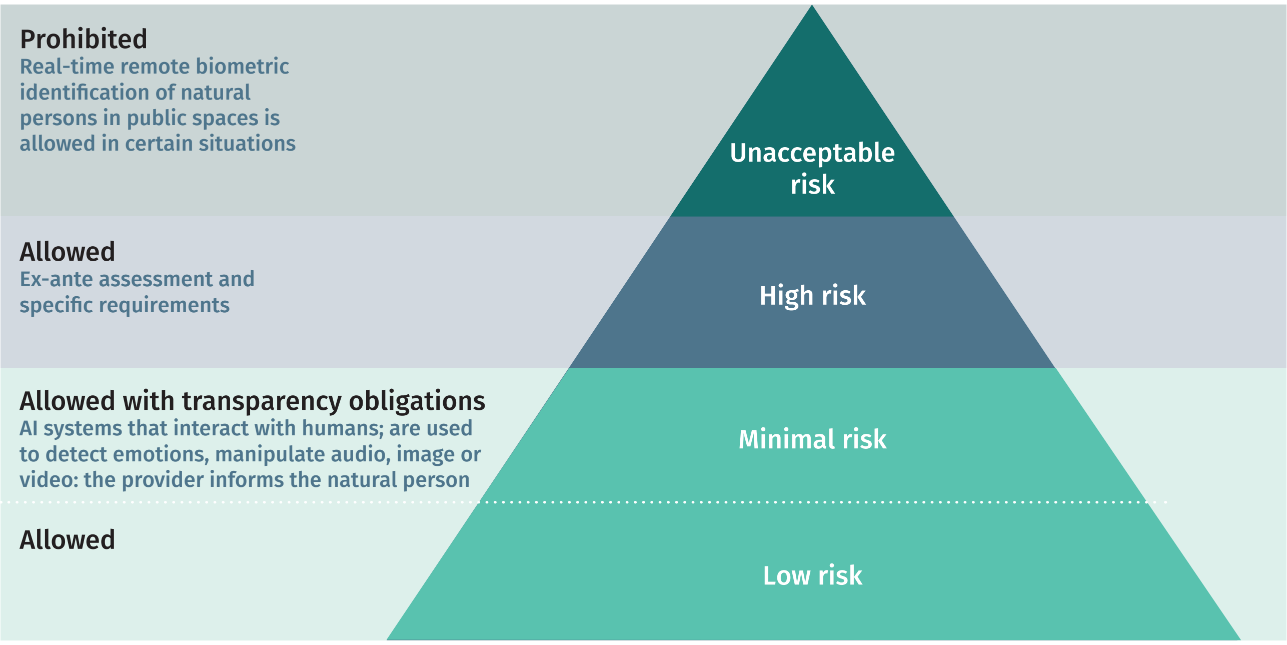
Users also have responsibilities and should use high-risk AI systems in accordance with the usage instructions. It is worth noting that the Regulation does not establish other direct obligations beyond what the provider estimates is necessary. Users must ensure that input data is relevant, monitor the operation of the system, keep logs and comply with their obligation to carry out a data protection impact assessment.

Governance, implementation and penalties

At EU level, a European Artificial Intelligence Board is to be created, of one representative for each national supervisory authority, together with the European Data Protection Supervisor, and chaired by the European Commission. Its tasks include collecting and sharing expertise, ensuring uniform administrative practices in member states and issuing recommendations on the implementation of the Regulation.

At national level, member states will designate one or more national competent authorities to supervise the application and implementation of the Regulation and lay down rules on penalties. Non-compliance with the rules related to prohibited AI practices or data and data governance will be subject to fines of up to €30 million or 6 per cent of total worldwide annual turnover.

Figure 1 **The AI Regulation risk-based approach**


Prohibited
Real-time remote biometric identification of natural persons in public spaces is allowed in certain situations
Unacceptable risk
Allowed
Ex-ante assessment and specific requirements
High risk
Allowed with transparency obligations
AI systems that interact with humans; are used to detect emotions, manipulate audio, image or video: the provider informs the natural person
Minimal risk
Allowed
Low risk

Source: author's elaboration.

Critical assessment

This assessment of the AI Regulation draws on the principles of the 'social shaping of technology' in which technological design and implementation are not just patterned by narrow technical considerations but also by a range of intertwined social, economic and political factors (Williams and Edge 1996). In such a process, many actors should intervene, including trade unions. Public engagement and diversified expertise are also crucial elements that contribute to a genuine 'social shaping' of technological change (Jasanoff 2003).

Uptake of AI

The recurrent narrative put forward by the EC is the need to establish ecosystems of trust and of excellence. The narrative runs thus: this Regulation addresses risks and, in so doing, creates trust; member states invest in AI and innovate and, in so doing, create excellence; and the result is an acceleration of the uptake of AI. Protecting fundamental rights, which should be the core objective, comes secondary. Protecting workers' rights is absent altogether.

The EC explains that its definition of AI, and the risk-based approach that underpins the draft Regulation, are based on public consultation. Reviewing the responses to its White Paper shows, however, that what is presented as a democratic process is far from it: an overwhelming majority of respondents were industrial actors and tech companies. They directed the regulatory process in a direction that serves their interests while pretending to serve knowledge, science and technology. The voice of society, in particular trade unions, was drowned in the noise.

A narrow approach

The precautionary principle should have been central to a Regulation whose subject matter is an uncertain and risky technology. Instead the EC has opted to use a narrow version of the risk-based approach: aspects such as the impact of AI on fundamental rights, workers' rights and the environment, as well as the need to anticipate how it will evolve in the future, are not sufficiently addressed.

This narrow approach prohibits a handful of unacceptable uses but allows all others. Low-risk uses are not defined and no differentiation is made between intermediary levels of risk within it; while uses in that category have free access to the market.

High-risk uses are allowed as long as the provider, based on self-assessment, is able to tick the boxes on what is a largely procedural checklist. The list of high-risk AI uses covers essential aspects of the life and work of European citizens (from education to employment and access to services, credit, justice, etc.) and little is being done to limit or prohibit them. Providers do not have to eliminate the risk, they are simply expected to notice it, monitor it and provide information about it. Why certain uses are placed in this category is not described and thus appears to be arbitrary.

Residual risks are deemed 'acceptable' as long as risk management measures are sufficient and the AI system is used in accordance with its intended purpose or under conditions of 'reasonably foreseeable misuse'. The only obligation is that those residual risks have to be communicated. This must be addressed as it leaves the door open to potential unknown dangers.

The list of prohibited uses is short, appears final and no provision is made in the Regulation to update and extend it.

Weak mandatory requirements

The requirements that providers of high-risk AI systems have to meet are weak and insufficient. In particular, self-assessment should be replaced, at the very least, by third-party conformity assessment in all cases and not take place only during system development. A harmonised assessment framework must be developed to avoid a multiplicity of approaches and, finally, where a user puts together several AI systems, the assessment should not only cover each individual system but also them all in combination.

When dealing with workplace introductions of AI systems, any assessment must be performed in full consultation with the social partners.

Furthermore the EC considers that standardisation should play a key role in helping providers comply with the Regulation. This raises two concerns: (a) the standardisation process is industry-led and the participation of societal stakeholders is not always guaranteed or possible; (b) standards should not be imported from other regional superpowers but developed in Europe. In the current equation, the end-user (citizen, worker, consumer, etc.) is absent.

Liability and absence of redress

The key issue of liability is not addressed by the Regulation. Who is responsible for the damage in which AI systems have played a role is a major issue. Given that they operate with a certain level of autonomy, the current liability approach is hard to implement and clarifying the relative liability of the provider vs. the user is essential.

Worse, another aspect omitted from the Regulation concerns redress. End-users who feel that they have been damaged by a particular AI system have no access to any specific mechanism for redress either against the users of such systems or the provider.

Questioning 'expertise'

AI is a complex field and the EC has relied on 'experts' to guide the regulatory process. It has consequently overlooked social and societal expertise, leaving technical experts to monopolise the regulatory frame. For example, the High-Level Working Group on AI tasked with developing ethical guidelines was industry-led. This bias is visible in how the Regulation describes the risk-based approach and the proposed form of governance via an AI Board. This important body should be more open and, while the Regulation states that it may invite 'external experts', the social partners are not specifically mentioned. Their presence would provide minimal compensation for the absence of an employment perspective across the text.

The need for an ad hoc directive on AI in employment

The principle of subordination in the relationship between employer and workers, and the prospect for the abuse of managerial power, justify the use of AI in the field of employment being taken out of the scope of the Regulation. Instead this should be regulated via a standalone ad hoc directive that focuses on protecting workers (including platform workers) and on enabling them to exercise their rights.

The AI Regulation lists as a high-risk use 'Employment, workers management and access to self-employment'. The systems listed include those used to recruit, select, advertise vacancies, screen or filter applications, evaluate candidates during interviews or tests, make decisions on promotion and termination of work contracts, allocate tasks, monitor, and evaluate performance and behaviour.

Every single one of these AI uses triggers risks whose severity is such that they deserve to be addressed through dedicated legislation. The rapid development of AI means that potentially every worker in every sector will be exposed to AI systems, with the data they generate being used by their employer and potentially shared with third parties in ways unknown to the worker.

Several recent cases of abuse have emerged with companies sentenced to large fines. H&M was fined €35 million in October 2020 for recording the private data of several hundred employees in a marketing service centre in Bavaria. After periods away from work, management would conduct so-called 'welcome back' meetings during which workers would be asked questions about their vacation, diseases, cancer treatment, etc. There were also 'flow talks' about family or religion. Management would then meticulously use this information to evaluate workers' performance and draw up detailed profiles used as the basis for later decisions on their employment. The German electronics retailer botebooksbilliger.de was fined €10 million in relation to its use of CCTV cameras to monitor employees and customers (CPDP 2021).

The ad hoc directive on AI in employment should address the following dimensions:

1. Employers' responsibilities in preventing AI risks. Risk is a function of exposure and hazard. In occupational health and safety, employers are obliged to conduct risk assessments. AI risks go beyond occupational health and safety as they include the possible abuses of managerial prerogative stemming from the nature of the employment relationship as well as other risks to privacy, fundamental rights, data protection and overall health. This is why a specific framework to assess AI risks in the context of employment, with the active involvement of workers and their representatives, is required. This should consider the different dimensions of AI systems (physical, behavioural, biological, emotional, neurological) and the intensity of workers' exposure to them. This would help in establishing preventive measures and in clarifying the liability of employers, developers and other users along the value chain.
2. Privacy and data protection rules. AI is data hungry and workers are an important source of personal data. Although GDPR is a powerful regulation, it is difficult in practice for workers to exercise their rights in front of an employer: to be informed; to gain access; to seek rectification; to erasure; to restrict processing; to data portability; to object; and to ask for the logic behind automated decision-making and profiling. The directive should also address other issues such as 'purpose limitation' (data should be used only for the purpose for which it is collected), 'profiling' and 'informed consent'. A framework to obtain valid informed consent (as an exceptional legal basis for the employer to process worker data) is required.
3. Algorithmic workplace explainability. Even more than transparency, what is needed is explainability. This means providing, in lay language, an understanding of how an algorithm has pushed, nudged or influenced matters in a certain direction and not simply a description of its inner workings. This is essential to any AI system that interacts with workers, particularly when it produces data whose potential uses are unknown. Explanations need to be adapted to the user profile (Adadi and Berrada 2018). A 'workplace explainability framework' should enable workers and their representatives to understand the role and impact of AI systems in the workplace and to avoid situations where they use AI passively and are mere data points or receivers. Workers should have agency over the AI

systems with which they interact and to be literate about them: to ask for explanations, understand how AI systems process inputs and deliver output and to contest them (GDPR Art. 22). Moreover, not all automated decisions should be permitted. The regulator should prohibit algorithms that can be used to terminate the employment relationship or are the de facto sole decision-maker about workers (Ponce del Castillo 2021).

4. The role of trade unions in ‘human-in-command’. The human-in-command principle should be present in all human-machine interactions as a new component of work organisation. This is understood as giving the last word to humans and explaining what data sources are responsible for final decisions where humans and machines are acting as an *assemblage* (Zanzotto 2019). Trade unions should be part of this and play an active role when they estimate that a human operator is needed. This principle can help anticipate, prevent and manage current and emerging AI risks.
5. Algorithmic worker surveillance. The monitoring of workers is regulated by national laws that often predate GDPR and do not cover modern and intrusive people analytics (Eurofound 2020; Moore 2020). Indeed, AI has brought this to a new level which we can define as algorithmic worker surveillance: advanced analytics (biometrics, machine learning, semantic analysis, sentiment analysis, Emotion Sensing Technology, etc.) can measure biology, behaviours and emotions. This is comparable to switching from radar, which scans the surface of the sea, to sonar, which builds a 3D image of everything under the surface. Algorithmic surveillance does not passively scan but ‘scrapes’ the personal lives of workers, actively builds an image and then makes decisions. The directive should prohibit algorithmic worker surveillance (Ponce del Castillo 2021).

The idea of a draft directive addressing AI in employment is already in circulation and the European Parliament, in its resolution on ‘A strong social Europe for Just Transitions’, has urged the Commission ‘to present a directive on minimum standards and conditions… to protect the health and safety of workers and to ensure… respect for working hours, leave, work-life balance and other digital rights at work such as the right to disconnect, the protection of workers’ privacy, including through remote monitoring or any other tracking, and the prohibition of microchip implants on workers and of the use of artificial intelligence in recruitment processes, while taking into consideration the European Social Partners Framework Agreement on Digitalisation’ (Radtke and Jongerius 2020).

Conclusion

The EC considers that its draft Regulation creates an ‘Ecosystem of Trust’, despite having stated in its White Paper that ‘lack of trust is a main factor holding back a broader uptake of AI’. This obsession with trust stems from a misrepresentation of what society, in particular workers and their trade unions, thinks about AI: it is not about being afraid or not trusting AI or any other technology but about objecting to specific uses that have been demonstrated to be excessive, disproportionate or which contravene fundamental rights. Those uses, in the specific context of employment, are compounded by the subordinate relationship between worker and employer.

The approach that underpins the Regulation, based on allowing high-risk uses as long as the provider carries out self-assessment and complies with certain requirements, deserves all the criticisms we make of it. The EC is giving technology providers priority with the objective of giving them the confidence to embrace AI and encourage businesses to develop uses.

Ultimately the EC is trying to position the EU as the leader of the global AI pack. Our concern is that it may be doing so at the expense of citizens and, in particular, workers. An ad hoc directive on AI in employment is thus absolutely necessary to re-balance the scales.

References

Adadi A. and Berrada M. (2018) Peeking inside the black-box: a survey on explainable artificial intelligence (XAI), IEEE Access, 6, 52138-52160.

CPDP (2021) 'Smile for the camera, you are being watched'. Workplace surveillance: enforcing workers' rights, paper presented at the Computers, Privacy and Data Protection conference, 3 March 2021. https://youtu.be/FJ8YRNSF-1k

Eurofound (2020) Employee monitoring and surveillance: the challenges of digitalisation, Luxembourg, Publications Office of the European Union. https://www.eurofound.europa.eu/publications/report/2020/employee-monitoring-and-surveillance-the-challenges-of-digitalisation

European Commission (2021) Proposal for a Regulation of the European Parliament and the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, COM(2021) 206 final, 21 April 2021. https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence-artificial-intelligence

Jasanoff S. (2003) (No?) accounting for expertise, Science and Public Policy, 30 (3),157-162.

Moore P. (2020) Data subjects, digital surveillance, AI and the future of work, Brussels, European Parliament. https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU(2020)656305

Ponce del Castillo A. (2021) Algorithmic workplace surveillance, paper presented at the Digit Debates Series, University of Sussex, 24 March 2021. https://youtu.be/udQyOmrJLQQ

Radtke D. and Jongerius A. (2020) Report on a strong social Europe for just transitions, Brussels, European Parliament. https://www.europarl.europa.eu/doceo/document/A-9-2020-0233_EN.html

Williams R. and Edge D. (1996) The social shaping of technology, Research Policy, 25 (6), 865-899.

Zanzotto F.M. (2019) Human-in-the-loop artificial intelligence, Journal of Artificial Intelligence Research, 64, 243-252

All links were checked on 2 June 2021.

ETUI publications are published to elicit comment and to encourage debate. The views expressed are those of the author(s) alone and do not necessarily represent the views of the ETUI nor those of the members of its general assembly.

The ETUI Policy Brief series is edited jointly by Nicola Countouris, Jan Drahokoupil, Philippe Pochet, Aída Ponce Del Castillo, Kurt Vandaele and Sigurt Vitols.
The editor responsible for this issue is Nicola Countouris, ncountouris@etui.org

This electronic publication, as well as previous issues of the *ETUI Policy Briefs*, is available at www.etui.org/publications. You may find further information on the ETUI at www.etui.org.

The ETUI is financially supported by the European Union. The European Union is not responsible for any use made of the information contained in this publication.