RESEARCH REPORT

Eurofound

Anticipating and managing the impact of change

Ethical digitalisation at work: From theory to practice

Ethical digitalisation at work: From theory to practice

Eurofound

European Foundation for the Improvement of Living and Working Conditions

When citing this report, please use the following wording:
Eurofound (2023), *Ethical digitalisation at work: From theory to practice*, Publications Office of the European Union, Luxembourg.

Authors: Sara Riso, Dragoș Adăscăliței and Ricardo Rodríguez Contreras

Research manager: Sara Riso

Research project: Human and ethical implications of automation and digitisation for work organisation and working conditions

Provider: This overview report is based on inputs from three different research strands coordinated by the VVA–Open Evidence consortium. These include expert views collected through two online Delphi studies; interviews with policy stakeholders representing governments, employer organisations and trade unions; and four case studies (conducted by Alba Boluda, Frans Folkvord, Karlijn Van Houten, Lasse Wulff Andersen, Laura López Forés and Lucie Lechardoy). Another three case studies were conducted by members of the Network of Eurofound Correspondents: Amanda Kinnunen and Vera Lindström (Oxford Research), Thilo Janssen (Institute of Economic and Social Research, Hans-Böckler-Stiftung) and Sandra Vogel (German Economic Institute).

Acknowledgements: The authors would like to thank the many experts who participated in the two Delphi studies. They also extend their gratitude to David Reichel of the European Union Agency for Fundamental Rights for reviewing the final draft of the report; to three experts – Adrián Todolí Signes, Andrea Bertolini and Kimberly Houser – who were consulted on this research and provided inputs to various research activities; and to Camille Guey, Laura Todaro and Lucie Lechardoy of the VVA–Open Evidence consortium, who coordinated the activities of their research teams.

Luxembourg: Publications Office of the European Union, 2023

Print: ISBN 978-92-897-2348-0 doi:10.2806/512736 TJ-07-23-333-EN-C
PDF: ISBN 978-92-897-2347-3 doi:10.2806/235924 TJ-07-23-333-EN-N

This report and any associated materials are available online at https://eurofound.link/ef22018

Research carried out prior to the UK's withdrawal from the European Union on 31 January 2020, and published subsequently, may include data relating to the 28 EU Member States. Following this date, research only takes into account the 27 EU Member States (EU28 minus the UK), unless specified otherwise.

The European Foundation for the Improvement of Living and Working Conditions (Eurofound) is a tripartite European Union Agency established in 1975. Its role is to provide knowledge in the area of social, employment and work-related policies according to Regulation (EU) 2019/127.

European Foundation for the Improvement of Living and Working Conditions

Telephone: (+353 1) 204 31 00
Email: information@eurofound.europa.eu
Web: https://www.eurofound.europa.eu

Contents

Executive summary

Introduction

Digitalisation is bringing about profound changes in how people live and work. While digital technologies can be harnessed for good and drive progress, they may also raise ethical concerns, depending on their field of application and their purpose.

The workplace is an important arena where technologies display their transformative potential but it often remains under the radar in policy debates on ethics and digital technologies. The implications of automation of work and digitisation of processes for working conditions are intertwined with ethical considerations. In data-driven, digitally connected and automated workplaces, new ethical issues arise. These relate, for example, to the capture and management of personal data, the transformation of the role of the human worker and changes to the nature of work, and distress or anxiety among workers who perceive their jobs as at risk.

Drawing on Delphi studies, in-depth interviews with policy stakeholders, and exploratory case studies, this report contributes to the broad discussion about the ethical implications of automation and digitisation technologies and the effects of such technologies on working conditions.

Policy context

In the race for technological innovation, the EU has taken a different approach from other global superpowers, such as China and the United States, placing ethics at the heart of technology development and use. The EU's vision, firmly anchored in fundamental human rights and democratic values, is of technology that works for society – 'technology for good purpose', rather than for social control or profit.

Several EU-level policy initiatives stress the importance of ethics in the development and use of artificial intelligence (AI). Documents that make such references include the European strategy on AI, the European Commission's communication on a coordinated plan on AI, and the *Ethics guidelines for trustworthy AI* produced by the High-Level Expert Group on Artificial Intelligence.

On the regulatory front, the EU has been steadily building on the existing legal framework for data protection. Through the forthcoming AI Act, which will affect AI use in the workplace, the EU intends to increase the safeguards in place to ensure a secure, human-centric and ethical technology-driven economy and society. The draft text classifies the use of AI systems for work management and recruitment as 'high risk' and hence subject to scrutiny, checks and conformity testing.

The rapid development of AI – albeit still largely confined to 'narrow AI' applications (which automate specific tasks) – has further intensified the EU-level policy debate about the ethical implications of digital technologies for work, employment and society at large. At European level, job quality issues associated with the digitalisation of work are increasingly debated from an ethical perspective. Such issues relate, for example, to mental strain due to the use of intrusive technologies to monitor work and the growth of an 'always on' culture, which can blur the boundaries between work and private life.

Key findings

- The concept of ethics is central to the adoption and use of technology in the workplace and foundational to quality of work. According to the experts consulted, the ethical use of technologies entails not only adherence to ethical principles and compliance with fundamental rights but also the anticipation and mitigation of negative impacts on working conditions. Their views suggest that the risks associated with technology implementation for working conditions, human rights and ethical principles should be assessed comprehensively as a first step in design and development.
- The concerns most commonly expressed by the policy stakeholders interviewed related to future skills needs. This issue has an ethical dimension: without adequate training provision and reskilling programmes, workers will be vulnerable to skills obsolescence and job loss. Other ethical concerns involved data protection and privacy, and trust – the latter in relation in particular to the transparency of AI systems and potential discriminatory outcomes.
- Social dialogue and collective bargaining can be slow to respond to challenges arising from the digitalisation of work. These may result from more intensive use of technologies for recruitment, work management, monitoring and surveillance, and profiling of workers. Nonetheless, some instances of collective bargaining outcomes demonstrate the added value of social dialogue in dealing with the digitalisation of work.

- Survey data and evidence from the case studies carried out for this report suggest that companies tend not to have explicit policies, guidelines or procedures addressing ethical issues arising from technology in the workplace. In the establishments interviewed, principles such as accountability, transparency and human oversight were loosely embedded in technical procedures or addressed in training. The experts consulted also indicated that ethical principles are not systematically embedded in the design and development of AI; instead, at the design stage a trial-and-error approach is often followed to identify ethical pitfalls of AI systems.
- Technology adoption in establishments was typically motivated by reducing costs and increasing productivity. However, the technologies were also intended to make work more human-centric and to improve working conditions by making work less repetitive and more rewarding. In several establishments, management regarded this as particularly important to retain staff in the face of increasing labour shortages or to align tasks with workers' qualifications.

Policy pointers

- There is a need to coordinate efforts among a range of different stakeholders to translate agreed ethical principles into tools and methods for designing and using technologies. Practical guidance on how to incorporate these systematically throughout the technology lifecycle must also be developed. In the establishments interviewed, future plans invariably involve greater automation of tasks, and this is expected to impact the quality of work to a much greater extent than seen and documented in previous research. There is a clear need for a more forward-looking and ethical approach to technological change.
- The effects on the workplace warrant more attention in the general policy debate on digitalisation and ethics. National policies focusing on ethics tend to be broad in scope and pay little attention to issues related to quality of work. Policy initiatives focusing on ethical workplace digitalisation should be incorporated into national digitalisation and AI strategies. Such initiatives might include campaigns raising awareness of the benefits of ethical technology design and use, efforts to embed ethics in education and training curricula or practical guidance for companies on implementing an ethical and human-centred approach to technology design and use.
- Greater efforts are needed to build the capacity of national social partners to deal effectively with issues arising from the digitalisation of work. The European social partners' 2020 framework agreement on digitalisation is an important instrument for coordinating such efforts and promoting the exchange of information with a view to mainstreaming digital ethics into collective bargaining and social dialogue. The involvement of social partners (beyond consultation) in the design and implementation of policy initiatives on ethical digitalisation and ethical AI must also be increased to create an increased focus on quality-of-work issues.

Introduction

The issue at a glance

At a time of rapidly increasing digitalisation, questions around the ethical and responsible design and use of technologies, particularly artificial intelligence (AI), have come into sharp focus in public and policy debates at both national and European levels. A string of EU policy and regulatory initiatives have given impetus to the strong emphasis on ethics in AI use in national debates (Eurofound, 2022a). These include the still intensely debated draft AI regulation, also known as the AI Act (European Commission, 2021a).

In the global race for technology leadership and innovation, the EU has taken an ethical stance on AI and on digitalisation more generally. This becomes particularly apparent when comparing the EU with other global leaders, such as the US and China, for which digital technologies and AI tend to be geared to objectives such as enhancing their military and economic power and their domestic surveillance capabilities. Through various items of legislation, the EU is seeking to ground digitalisation in respect for fundamental human rights and democratic values.

However, technologies are moving targets. This makes it difficult to get to grips with the ethical and regulatory challenges that digitalisation poses. The workplace is an important arena in which technologies display their transformative potential and, in doing so, raise a host of ethical issues. By accelerating technology adoption across many industries, the COVID-19 pandemic also brought to light some of the ethical pitfalls of the use of new workplace technologies (Eurofound, 2022a).

Examples of concerns of an ethical nature that became more prominent during the pandemic relate to the expanded monitoring and surveillance capabilities of digital technologies and the wide-ranging implications for working conditions and fundamental rights (Eurofound, 2020a). The rights to privacy and data protection are often threatened by intrusive digitally enabled workplace monitoring and surveillance. Other fundamental rights are also at stake. Intrusive digital surveillance at work is a threat to the right to freedom of association and collective bargaining, as it weakens workers' negotiating and organising power and heightens the imbalance of power in employment relationships. In addition, the invisibility of the control and supervision carried out by digital monitoring or surveillance systems makes it difficult for employees to contest management or employment decisions based on the data collected. This could potentially curtail workers' right to seek redress or remedy where decisions solely or heavily based on the data collected by such systems are unfavourable for them or perceived as unfair or discriminatory.

Concerns are also mounting around other consequences of the increased digitalisation of work. These concerns relate to the extent to which digital technologies transform the human role in the workplace, redefine the meaning of work, lead to a loss of acquired skills among workers or instil fear of future job loss, which may compromise workers' well-being and generate anxiety. The implications of increased digitalisation for quality of work have not only an ethical dimension but also legal ramifications, going well beyond the potential infringement of privacy and data protection rights and extending, for example, to the rights to human dignity, integrity of the person, and fair and just working conditions, as enshrined in the Charter of Fundamental Rights of the European Union (CFREU).[1]

Other controversial uses of technologies relate to the use of algorithms for work management and recruitment software, which comes with its own set of ethical risks, including algorithmic bias, loss of privacy, power asymmetry, a lack of transparency, unclear accountability and potential loss of human oversight. According to the draft AI Act, the use of AI in the recruitment and management of workers is currently considered 'high risk', and therefore requires companies to ensure human oversight and proof of safety and efficacy.

Principles such as trustworthiness, transparency, accountability and responsibility, justice and fairness are often considered by ethicists with regard to the use of AI in particular and are central to many ethical guidelines on it, including the European Commission's *Ethics guidelines for trustworthy AI* (AI HLEG, 2019). Some of these principles have a legal basis and are, for example, embedded in the EU General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679). Although the GDPR does not mention any specific digital technologies, the core data protection principles[2] have a significant impact on the way data-driven technologies are used in the workplace. In the absence

1 The charter became legally binding in the EU with the entry into force of the Lisbon Treaty in December 2009. The charter has six main titles under which relevant rights are grouped: dignity, freedoms, equality, solidarity, citizens' rights and justice. Information on the charter is available from an online resource of the European Union Agency for Fundamental Rights at https://fra.europa.eu/en/eu-charter

2 The seven data protection principles in the GDPR are (1) lawfulness, fairness and transparency; (2) purpose limitation; (3) data minimisation; (4) accuracy; (5) storage limitation; (6) integrity and confidentiality (security); and (7) accountability.

of clear and explicit data governance, the most intrusive, pervasive and opaque technologies – for example, the internet of things (IoT) and AI systems – could potentially violate the GDPR.

As digital devices become more affordable and their computing power increases, digital technologies will continue to penetrate the workplace at an increasing rate and will become more sophisticated and powerful. With power, however, come choices and responsibilities for all stakeholders, from employers and employee representatives in companies to policymakers at all levels. As technologies become more powerful (and more central to working life), worker protection must also evolve to address the multifarious ethical challenges posed by digitalisation. This should be part of a much broader discussion and rethink of the social contract for work, to ensure that technologies are harnessed for positive purposes and without compromising hard-won workers' rights.

In the context of a continuously evolving background, this report brings the focus of the discussion around the ethics of digital technologies back to the workplace, which has thus far received less attention in both public and policy debates on ethics. The report is intended to feed into the debate on the (good) use of technologies to benefit both workers and employers, and in the dual interest of safeguarding workers' fundamental rights and quality of work and enhancing productivity. These are fundamental ethical issues that should be investigated and debated at a time when social partners and public authorities are in a position to influence and shape future developments.

The report is for European and national policymakers, especially those who are engaged in discussions on the future of work or who are responsible for designing and implementing digitalisation and AI policies. It also addresses social partners, who are confronted with new ethical challenges arising from the ever-increasing digitalisation of work and who are called on to find common solutions in the interests of safeguarding workers' well-being and enhancing productivity.

Terminology and definitions

Digitalisation, automation and digitisation

Digitalisation is a broad term that refers to the increasingly widespread adoption and use of digital technologies with transformative effects on employment and on society more generally (Eurofound, 2018a). Key driving forces – or vectors of change – of this digital transformation are the automation of work and the digitisation of processes (Table 1). These are expected to have a fundamental impact on work and employment.

Table 1: Definitions of vectors of change and examples of associated technologies

Vector of change	Definition	Examples of associated technologies and applications
Automation of work	The replacement of (human) labour input with machine input for some types of tasks within production and distribution processes.	Advanced robots are endowed with sensors and high-level, dynamic programming, enabling them to perform tasks that require more flexibility and accuracy than those performed by traditional industrial robots. One type of advanced robot is collaborative robots or cobots (used either as a service robot or in manufacturing). These are equipped with software-controlled sensors and collaborate with humans in a shared work environment. Other types are advanced automated guided vehicles and mobile robots with dynamic algorithmic routing capabilities. Such robots are often used in business logistics for transportation tasks. AI capabilities may be embedded in a robot frame capable of motion or may be contained within an existing computing infrastructure (including chatbots). Robotic process automation involves robotic software handling structured (or semi-structured) data and performing manual, repetitive and rule-based tasks and processes. Combined with AI, robotic process automation software can perform tasks involving greater cognitive ability and handle complex processes requiring learning.
Digitisation of processes	The use of sensors and rendering devices to translate (parts of) the physical production process into digital information (and vice versa), and thus take advantage of greatly enhanced possibilities for the processing, storage and communication of digital information. While sensor technologies translate physical information (for example, temperature, sound and light) into digital information, rendering devices do the opposite, translating digital information into analogue information (for example, using a printer).	IoT systems or applications comprise interconnected sets of entities, including miniaturised computing devices, sensors, and components, and collecting and processing large amounts of data. The internet is the channel through which data flow and through which objects are connected. The data processing capabilities of IoT systems can be further boosted by the use of AI. IoT and AI are expected to converge to form what has been labelled as the artificial intelligence of things. Wearables are small mobile computers incorporating electronics, software and sensors that are connected to the cloud. They can also be considered part of the IoT. Worn directly on the body, wearables record employees' activities using sensors.

Notes: *Another vector of change identified in Eurofound's conceptual framework is 'coordination by platforms', which is not listed here as it is outside the scope of this study. The list of technologies provided in the table is not exhaustive and is restricted to those of ethical relevance.*
Source: *Eurofound, 2018a, 2018b, 2019, 2020b; Forbes, 2019*

It should, however, be noted that automation and digitisation technologies are interconnected and synergetic. For example, advanced robots or machines are often equipped with IoT or other sensor technologies that read and, in some cases, learn from the environment; they are increasingly connected to the cloud, enabling them to share and analyse data that feed into AI systems.

Artificial intelligence

At present, there is no single agreed definition of AI. Rather than a single technology, AI refers to a collection of different techniques used to solve problems that would otherwise require human intervention. AI can power other automation and digitisation technologies, such as the IoT and robotics, and can expand their capabilities, further reducing the need for direct human participation.

A distinction is often made between general or strong AI and narrow or weak AI. Narrow or weak AI can automate specific tasks, while general or strong AI is capable of replicating human thinking and learning. The latter has been only theoretically possible until recently, popularised by science fiction movies depicting a techno-dystopian, AI-driven future, such as *Blade Runner*; *I, Robot*; and *I am Mother*.

Far from science fiction, AI is most often associated with machine learning, which is one of the most used, albeit not yet mainstream, applications of (narrow) AI and 'the main driver of the current AI wave' (Stahl, 2021). Machine learning refers to techniques that use algorithms to harvest and analyse vast amounts of data to make predictions or decisions without explicit programming. These algorithms are capable of self-learning and self-optimisation. One of these techniques is deep learning, which involves the use of complex algorithms and deep neural networks to train a model. A subset of deep learning is generative AI, of which the most well-known application is ChatGPT.

In the first draft of the proposed AI Act, an AI system is defined as 'software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with' (European Commission, 2021a, p. 39). The range of approaches and techniques that, according to Annex I, are included in the above definition are:

> *(a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;*
>
> *(b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;*
>
> *(c) Statistical approaches, Bayesian estimation, search and optimization methods.*

This list of approaches and techniques is intended to be kept up to date (recital 6 of the draft AI Act).

After several revisions, the most recent compromise text approved by the Council of the EU (2022a) narrowed the definition to 'systems developed through machine learning techniques and knowledge-based approaches'. In the revised text, the category 'general purpose AI systems' is included. These can be used for various purposes and a broad range of tasks. They are defined as systems 'intended by the provider to perform generally applicable functions such as image and speech recognition, audio and video generation, pattern detection, question answering, translation and others' (Council of the EU, 2022b).

Research methods

This report summaries the key findings of a study combining different and complementary research methods. This mixed methods approach was applied owing to the complexity as well as the novelty of the issue under investigation.

Delphi studies

One research strand involved two Delphi studies, each comprising two survey rounds conducted online. The Delphi technique is a widely recognised research method that brings together expert opinions through a series of iterative questionnaires, with a view to coming as close as possible to a group consensus. The Delphi method has also been extensively used as a tool (known as ethical Delphi) for eliciting and exchanging views and opinions between experts on ethical issues in different areas (Hansson, 2017).

For both Delphi studies, consensus was defined as the agreement of a substantial majority (namely 75% or higher) of Delphi panellists over the proposed statements and items covered in the questionnaires. In some instances, the statements proposed in the first survey round were modified based on the input received from the experts consulted and reproposed in the second round. The experts also provided commentaries on the different statements and items they were asked to rate, using a five-point Likert scale.

Delphi study on ethics in the context of automation and digitisation

The first Delphi study explored the meaning of ethics in relation to automation and digitisation in the workplace and the ethical implications of the use of AI, advanced robotics and IoT technologies for work. The views collected fed into a broad conceptualisation of ethics in relation to the use of digital technologies in the workplace. A caveat of this exercise is that the ethical appraisal of the implications of the use of new or emerging technologies is inevitably incomplete, as the effects are yet to be seen. In this Delphi study, the statements proposed to the experts consulted were formulated largely based on previous desk research compiled in a separate publication (Eurofound, 2022a).

The expert panel consisted of academics and scholars with expertise in workplace and ethical issues related to the specified technologies. Individuals with legal and non-legal backgrounds were included in the expert panel. Some 21 experts (11 women and 10 men) participated in the first survey round and 16 in the second and final round (9 women and 7 men), out of a total of 83 experts invited to participate. Consensus was reached on 27 of the 33 proposed statements and items in the first round, and 19 of the 27 statements and items in the second and final round.

Delphi study on ethical design and development of machine learning applications

Narrowing the focus of the research to machine learning – which is one of the most established examples of narrow AI – the second Delphi study explored the importance and embeddedness of ethical principles in technology design and development, characteristics of machine learning that raise ethical concerns, and key requirements or pre-conditions for the ethical use of machine learning in the workplace.

Twenty-three technology experts (13 women and 10 men) – with technical knowledge in AI – participated in the first survey round and 16 (9 women and 7 men) in the second and final round, out of a total of 200 experts invited to participate. Consensus was reached on 16 of the 18 proposed statements and items in the first round, and 12 of the 16 statements and items in the second round. Both survey rounds included open-ended questions to collect further insights into the statements and items rated.

Interviews with policy stakeholders

Another strand of research consisted of 58 in-depth qualitative interviews with policy stakeholders, representatives of governments, trade unions and employer organisations in nine EU Member States, namely Belgium, Denmark, Finland, France, Germany, the Netherlands, Poland, Spain and Sweden (Table 2). Apart from Poland, these are countries where, according to Eurofound's previous research (Eurofound, 2022a), the debate on the digitalisation of work is most prominent and long standing. The interviews were instrumental in collecting the views of stakeholders on pressing ethical issues related to technology use and supplementing and updating the information previously collected through the Network of Eurofound Correspondents, in relation to debates, policies and regulatory developments.

Table 2: Number of stakeholder interviews, by country and type

Country	Governments	Employer organisations	Trade unions	Total
Belgium	2	2	2	6
Denmark	3	1	2	6
Finland	1	2	2	5
France	2	2	3	7
Germany	2	3	3	8
Netherlands	3	2	2	7
Poland	2	2	2	6
Spain	3	2	2	7
Sweden	2	2	2	6
Total	**20**	**18**	**20**	**58**

Source: *Authors' elaboration*

Exploratory case studies

The last strand of research involved the analysis of seven exploratory case studies of establishments in selected EU countries. These investigated a variety of approaches taken in these establishments to adopting and implementing digital technologies in work processes, if any ethical issues emerged in the process, and how these were dealt with by management.
The case studies mainly drew on interviews with innovation and line managers, employee representatives, and employees using the technology on a regular basis. This information was supplemented by documentation from company web pages and internal documents made available by the establishments interviewed.

The interviews were conducted in accordance with a semi-structured interview guide, designed around a set of core themes and questions intended to explore the implications of technology adoption and use for ethics and working conditions.

As the cases venture into the uncharted territory of ethical digitalisation, they are exploratory in nature.[3] They are by no means representative of all establishments, including those of the same size or operating in the same sector or country. Table 3 provides basic information about the cases investigated, illustrating their variety in terms of the technologies they cover and the characteristics of the establishments they involve (including the forms of employee representation).

Table 3: Overview of exploratory case studies

Case study example (location)	Sector (Nomenclature of Economic Activities)	Establishment size (number of employees)	Ownership structure	Form(s) of employee representation	Technology in focus
Lynred (Veurey-Voroize, France)	C26.11 – Manufacture of electronic components	1,037	Owned by Thales (50%) and Safran Electronics & Defense (50%)	Works council	AI-based solutions for product quality control and automated dispatching of helpdesk request tickets to the information technology department
Container Terminal Altenwerder (Hamburg, Germany)	H50.20 – Sea and coastal freight water transport	550	Owned by public limited companies HHLA (75%) and Hapag-Lloyd (25%)	Works council and trade union	Technologically advanced automated guided vehicles and other advanced automated systems
Municipality of Södertälje (Sweden)	O84.11 – General public administration activities	6,500 (of whom 21 work at the contact centre and 18 work at the payroll office)	Public entity	Trade union	Robotic process automation in the payroll office and AI chatbot in the call centre unit
Manufacturing company (Finland)*	C26 – Medical equipment manufacturing and pharmaceutical manufacturing	656	Finnish subsidiary of a multinational conglomerate	Trade union and occupational safety and health representation	Collaborative robots (cobots) on assembly lines
Orange España (Madrid, Spain)	J61 – Telecommunications	2,065	Spanish subsidiary of French multinational company Orange	Works council and trade unions	Robotic process automation widely applied from human resources (HR) to customer care and finance functions
Wicro Plastics (Kessel, the Netherlands)	C22.2 – Manufacture of plastic products	150	Stand-alone limited company	Works council	Advanced manufacturing robots enabling unmanned production on assembly lines
Civiløkonomernes A-kasse ** (Copenhagen, Denmark)	O84.3 – Compulsory social security schemes	100 (of whom 25 to 30 are careers advisers)	Non-profit organisation	Informal worker representation (staff committee)	Machine learning-based system for screening job applications and providing routine feedback (referred to as 'instant feedback')

Notes: ** Case has been anonymised at the request of the company. ** Civiløkonomernes A-kasse is a state-recognised unemployment funding association that compensates members financially for loss of income should they become unemployed. Similar insurance systems exist in only Finland and Sweden.*
Source: *Eurofound, 2021–2022*

3 Yin (2018) makes a distinction between three key types of case studies: exploratory, explanatory and descriptive. Exploratory case studies are considered an introduction to social research in areas where little is known about certain topics.

Structure and scope of the report

Building on a previous analysis of the ethical implications of technology use in the workplace (Eurofound, 2022a), this report investigates in greater depth the ethical considerations relating to workplace technologies from a theoretical and a more practical perspective. Although fraught with ethical challenges, technology use in the context of platform work and various forms of self-employment or freelancing is outside the scope of this report. The emphasis is on workplace challenges and practices in relation to technology design, adoption and use.

Drawing on the views of the experts consulted through two online Delphi studies, Chapter 1 explores the meaning of ethics and ethical implications of technology use in a digital workplace. It also discusses the role of technology design – particularly in relation to machine learning – in deploying technologies in a responsible, human-centric and ethical way.

Chapter 2 summarises the views and perspectives of the stakeholders consulted (in the selected EU Member States) on ethical issues that feature most prominently in national policy debates on digitalisation. It also provides an update on policy initiatives relevant to digitalisation initially mapped out in a previous report on ethics in the digital workplace (Eurofound, 2022a).

Through an analysis of selected case studies, complemented by desk research, Chapter 3 reviews the impact of selected automation technologies – including AI technologies – on working conditions from an ethical angle and describes management and organisational practices that can be used to ensure a human-centric and ethical approach to technological change in the workplace. These case studies are not in themselves examples of best practices; they only exemplify a variety of approaches to dealing with technological change in the workplace and the extent to which ethical issues are considered.

The report ends with conclusions and policy pointers derived from an analysis of the information collected during the study and the feedback provided by the experts and stakeholders consulted.

1 Meaning of ethics in a digital workplace

The term 'ethics' deserves some preliminary explanation before delving into an analysis of the ethical implications of digital technologies in the workplace. At its core, 'ethics' can be defined as what is morally right; the term commonly refers to rules and principles guiding human behaviour and actions and intended to produce positive outcomes. Such rules and principles clarify the obligations and responsibilities of the people concerned.

As technologies evolve and become more powerful, the connection between technology and ethics becomes more apparent. With technologies profoundly changing ways of working and creating new possibilities, it is increasingly important to understand and recognise the ethical implications arising to stop 'the deflection of human responsibility in technology-instrumented activities, especially when something goes wrong' (Johnson, 2005, p.4).

The concept of digital ethics (Capurro, 2017) – which is related to standards and principles applicable to a technologically driven society – is becoming increasingly central in the public debate on ethical, responsible and socially acceptable digitalisation. Unlike the public debate around the ethics of digitalisation, which has gained traction in recent times, the scholarly debate is a long standing one, with ethicists seeking to catch up with novel issues arising from the use of new and continuously evolving technologies. In the field of computer ethics, Johnson (1985) refers to 'new versions of standard moral problems and moral dilemmas, exacerbating the old problems, and forcing us to apply ordinary moral norms in uncharted realms' (p. 1).

Note on the main normative approaches to ethics

The concept of ethics is strongly underpinned by moral philosophy. A tenet of moral philosophy is that moral thinking needs to be supported by a moral theory that encapsulates the principles on which ethical judgements are based ('normative ethics'). In moral philosophy, a broad distinction is generally made between consequentialist and deontological theories, of which there are many varieties. One main difference between consequentialism and deontology is that the former assesses the moral quality of an action based on its outcomes, while the latter does not link the assessment of the moral quality of actions to outcomes.

Utilitarianism – one of the most prominent forms of consequentialism – advocates for ethical decisions to promote the greater good, rather than individual well-being. An implication is regarded as positive if it promotes the general welfare more than any other alternative (Shaw, 2017). However, utilitarianism faces both a measurement problem (how to measure utility) and a conceptual problem (how to define utility for a society) when used to understand the ethical dilemmas raised by powerful technologies such as AI.

In deontological theories, morality is grounded in a set of duties and obligations, and the ethical quality of an action lies in the principles or norms on which an agent acts rather than the effects of those actions on others or consequences for human welfare (Waller, 2011). Many ethical guidelines related to AI can be ascribed to a deontological understanding of ethics, setting universal principles guiding the ethical development, deployment and governance of AI.

Besides consequentialist and deontological theories, another normative approach to ethics, which is often discussed in relation to technologies, is virtue ethics. This approach focuses on the pursuit of goodness and assigns importance to guidance for individuals to help them to develop a virtuous character as a compass for ethical behaviour and judgements. This approach has informed ethical considerations in the area of information technology (IT) (Stahl, 2021) and laid the foundations for what has been labelled 'flourishing ethics' (Bynum, 2006). It gives importance to human values and principles – for example, freedom, autonomy, equality, justice and benevolence – that are conducive to humans flourishing in a technological age. The underlying idea is that technologies are to be developed and used to promote human flourishing or for social purposes (Powell et al, 2022). This is very much in line with the goals that the European Commission has set for Europe with regard to AI development and use (see European Commission, 2021b). The notion of 'human flourishing' also recurs in the European Commission's *Ethics guidelines for trustworthy AI*, developed by the High-Level Expert Group on AI (AI HLEG).

> *AI is not an end in itself, but rather a promising means to increase human flourishing, thereby enhancing individual and societal well-being and the common good, as well as bringing progress and innovation.*
>
> (AI HLEG, 2019, p. 4)

Reframing ethics in the context of workplace automation and digitisation

Ethics is an all-encompassing and multidimensional concept in the context of the automation and digitisation of work. There was consensus among the experts consulted in the first Delphi study that the deployment of automation and digitisation technologies has profound ethical implications in several areas of working conditions and may simultaneously involve a host of fundamental rights and other ethical principles. The ethical use of such technologies in the workplace therefore requires adherence to a wide range of ethical principles and compliance with fundamental rights, as well as the anticipation and mitigation of negative impacts on working conditions. Taken together, the expert views converge as to the importance of ethics for good-quality jobs in an increasingly digitised and automated workplace.

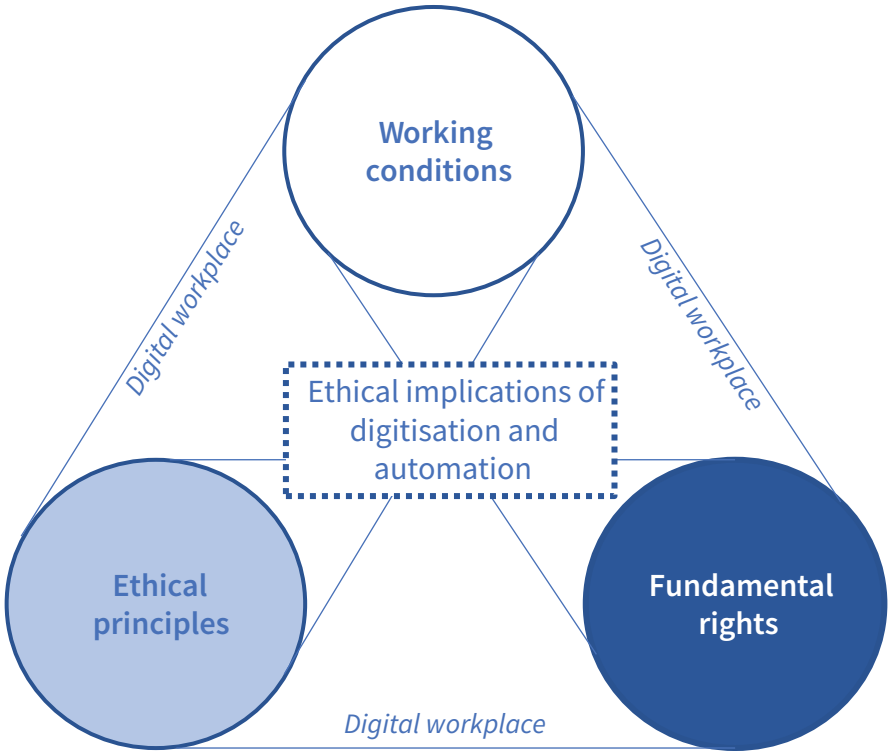
For the purpose of the analysis, implications for working conditions were grouped in the first Delphi study into the following categories, which are based on Eurofound's job quality framework (Eurofound, 2017):

- intrinsic quality of work (autonomy, skills and social support)
- health and safety (workplace-based physical and psychosocial risks)
- working time and work–life balance (duration, scheduling, flexibility and work intensity)
- employment quality (development opportunities and contractual stability)

Ethical issues connected to changes in work content and methods were also identified and regarded as intertwined with other aspects of working conditions, particularly autonomy and skills.

Intrinsic quality of work was regarded by the experts consulted as more inherently connected to the concept of work ethics and values (Figure 1).[4] For example, workers' perception of the value of their work and their commitment to work may be influenced by the degree to which they have autonomy or learn new things in their job and whether the social environment is stimulating and enriching for them.

Figure 1: Conceptualisation of the ethical implications of automation and digitisation technologies in the workplace



Source: *Authors' elaboration, based on the results of the first Delphi study, 2022*

4 Work ethics are underpinned by personal and work values. Personal values are related to the centrality of work in life and workers' moral requirement to perform work conscientiously. Work values can be intrinsic, extrinsic or social. Intrinsic values relate, for example, to workers' pursuit of autonomy and interest in personal development and creativity at work; extrinsic or material values are connected to achieving job security and a good income; and social values relate to the recognition of work as instrumental in securing positive social relations.

Ethically sensitive application of technology

Besides AI-powered algorithmic work management technology, with monitoring and surveillance capabilities, other applications of 'narrow' AI were identified by the experts consulted as posing ethical challenges. These include smart assistants for work management, emotion and facial recognition software and AI-based applications for assisting HR recruitment processes, with varying levels of human oversight, people analytics and automated profiling (for example, to predict whether employees will meet their targets). They also include clinical applications in areas such as imaging and diagnostics and health apps that can assess symptoms and recommend treatments.

The use of advanced robotics in the workplace – especially cobots and AI-enabled robots – is also regarded as having wide-ranging ethical implications for work. In an automated environment, more complex interactions with technology such as advanced machines are expected to deeply transform workers' roles. As for digitisation, the IoT raises great ethical concerns due to its tracking and monitoring capabilities, particularly when the technology is used for surveillance purposes and performance monitoring. IoT and sensor technologies – for example, incorporated in wearables – have greatly expanded the possibilities for collecting and analysing data and have extended employee monitoring well beyond task performance, to include workers' feelings, physiology, location and movements. As pointed out by scholars (see, for example, Wachter, 2018), the inherent and defining characteristics of the IoT – that is, the pervasive collection and linkage of user data – result in significant legal challenges.

Characteristics of machine learning relevant to ethics

Narrowing the focus to machine learning applications, a follow-up Delphi study conducted with technology experts sought their views on the characteristics of applications of ethical relevance. These included the opacity and unpredictability of outcomes; replicability of outcomes and accountability for them; data quality; and hidden or unnoticed biases.

These characteristics are in many ways interrelated. For example, the opacity and unpredictability of outcomes are connected, as machine learning applications are intrinsically complex. The issue of opacity is often referred to as the 'black box' problem, which particularly concerns non-symbolic AI techniques such as machine learning.[5] Even designers or developers may not fully grasp the internal workings of such applications, which by nature are dynamic due to their inherent capacity for self-learning.

The unpredictability of outcomes is also linked to the issue of the replicability of outcomes, as different inputs and parameters may lead to different outcomes every time the algorithm is run. However, from a regulatory perspective, it is essential to open the black box and understand the root causes of the unpredictability of machine learning systems to be able to determine liabilities for any harm or damage caused to those affected by the algorithmic decisions.

Finally, data quality is an overarching characteristic of ethical (and legal) relevance; machine learning applications create a need to collect as much data as possible, leading to the collection and use of unvetted or low-quality data. The excessive collection of data may also go against the principle of data minimisation in the GDPR. Furthermore, the assumption that collecting large amounts of data is enough to remove potential biases is flawed. Reflecting on overestimations in flu predictions in Google Flu Trends, Lazar et al (2014) highlight the importance of 'measurement and construct validity and reliability and dependencies among data' regardless of the volume of the data collected and used.

Ethical risks of workplace automation and digitisation

Fundamental rights

According to the experts consulted, the deployment of AI, IoT and advanced robotics technologies in the workplace has potential implications for fundamental rights, particularly the rights to human dignity (Article 1 CFREU), integrity of the person (Article 3 CFREU) and data protection and privacy (GDPR; Articles 7 and 8 CFREU). The use of AI and IoT technologies is expected to involve a broader range of fundamental rights. According to research on the impact of AI on fundamental rights conducted by the European Union Agency for Fundamental Rights (FRA, 2020), the use of AI may affect a very wide variety of fundamental rights, depending on the technology and specific area of use.

Table 4 lists the fundamental rights regarded by the experts as potentially at risk in the context of the automation of tasks and the digitisation of work processes. For most items, expert consensus was reached in the first survey round. The right to freedom of expression in relation to AI and IoT use was identified by the experts in the first round and consensus reached in the second round.

5 Non-symbolic AI techniques are statistical methods involving the execution of calculations to solve specified problems. They do so without being provided with real-world symbolic representations and have instead to create such representations based on large amounts of data. One of the drawbacks is that the generated representations are mathematically abstract and difficult to understand (see explanation in DataDrivenInvestor, 2018).

Table 4: Fundamental rights most at risk when using AI, advanced robotics and IoT in the workplace

	AI	IoT	Advanced robotics
Right to human dignity (Article 1 CFREU)	✓	✓	✓
Right to integrity of the person (Article 3 CFREU)	✓	✓	✓
Right to liberty and security (Article 6 CFREU)	✓	✓	
Right to data protection and privacy (GDPR; Articles 7 and 8 CFREU)	✓	✓	✓
Right to freedom of expression (Article 11 CFREU)	✓	✓	
Right to freedom of association and collective bargaining (Articles 12 and 28 CFREU)	✓	✓	
Right to non-discrimination (Article 21 CFREU)	✓		

Source: *First Delphi study, 2021*

To varying extents, the three technologies can also affect the right to fair and just working conditions, which provides that 'every worker has the right to working conditions which respect his or her health, safety and dignity' and 'every worker has the right to limitation of maximum working hours, to daily and weekly rest periods and to an annual period of paid leave' (Article 31 CFREU). Like other fundamental rights, this right is granted additional protection through secondary legislation, namely the European Working Time Directive (Directive 2003/88/EC) and the EU Framework Directive for occupational safety and health (Directive 89/391/EEC).

The views of the experts on issues covered under the right to fair and just working conditions were sought as part of a general assessment on the ethical implications of digital technologies for working conditions (presented in the section 'Working conditions').

Right to human dignity

The use of automation and digitisation technologies in the workplace can limit human dignity, which is foundational to ethics and, more generally, fundamental rights. The right to human dignity has a broad scope, which is not limited to the condemnation of extreme forms of violation or abuse but more generally protects the ability of human beings to be in control of their actions.

The right to human dignity is very relevant in the context of the digitalisation of work, as many digital technologies may influence workers' ability to act as purposive agents. According to the experts consulted, while the automation and digitisation of mundane and repetitive work tasks may be desirable for many, automating and digitising complex and meaningful tasks may deprive workers of experiencing the intrinsic value of performing their jobs and potentially lead to the dehumanisation of work, with an adverse impact on workers' well-being, job satisfaction and self-esteem. In order to preserve a real sense of agency, it is therefore important that technologies are deployed to augment human capacities and make jobs more rewarding and fulfilling.

Right to integrity of the person

The right to integrity of the person (with specific reference to mental integrity) was also considered at stake by the experts consulted in situations where AI and IoT technologies are used to monitor workers and track their activities, thus leading to 'surveillance anxiety' and stress. In the case of advanced robotics, working at a speed and pace determined by machines (for example, in human–robot collaborative systems) is an additional stressor with potentially detrimental psychological effects.

Right to privacy and data protection

Due to their extensive data gathering and processing capabilities, digital technologies have the potential to interfere with the rights to privacy and data protection. These are two separate but interconnected rights, with the difference being that the right to privacy is a universal right while the right to data protection does not (yet) have such status (EDPS, undated).

The risk of breaches of privacy and data protection rights are heightened in situations where AI and IoT-enabled technologies are used to track workers' performance, behaviours, location and movements. In relation to advanced robotics, the collection of personal information is inevitable and often a by-product of automated production processes. This, however, does not alleviate concerns around the legitimacy of monitoring workers.

The risk of breaches of privacy and data protection rights becomes even more acute in remote and hybrid working, where workers use digital devices for both personal and work reasons. This leads to the increased enmeshing of workers' private and working lives and the merging of personal with work-related data. There also exists the potential for hackers to access digital devices and get hold of workers' data, further intensifying concerns around privacy and data protection.

Rights to liberty and security, freedom of association and collective bargaining, and freedom of expression

Depending on the fields of application, both AI and IoT technologies can have an impact on other rights, namely the right to liberty and security, the right to freedom of association and collective bargaining, and the right to freedom of expression. According to the experts consulted, these rights can be potentially breached particularly in situations where the technologies feed into each other and are used for work management purposes. Such use can limit workers' ability to take decisions autonomously and alter power structures within the workplace to the advantage of employers, who have access to detailed data about workers that can be used to increase managerial control.

The use of AI may impose further limitations on the exercise of the right to freedom of expression outside the workplace. While workers can still communicate freely through public forums or social media, this freedom can be de facto compromised by the use of AI-based and automated recruiting and profiling software that extracts and scans information from social media accounts. As noted by some experts, the use of AI tools for HR purposes may also limit freedom of expression, as workers may be afraid that their social media activities are being monitored by their current or prospective employer.

According to the experts, several factors determine the extent to which digital technologies affect workers' freedom of expression and their organising power. These factors relate, for example, to the purpose of the technology being used, the presence of clear data governance policies, the configuration of the monitoring itself (if the technology is used for this purpose) and the quality of feedback from management.

Right to non-discrimination

Much has been said in the public and scholarly debates about the potentially discriminatory outcomes of AI algorithms due to societal biases entrenched in the raw data being used (for example, about gender, age and race). Several reports by the European Union Agency for Fundamental Rights have explored issues related to the right to non-discrimination (FRA, 2018, 2020), including data quality (FRA, 2019) and biases in algorithms (FRA, 2022).

Popular media has contributed to bringing the risk of algorithmic discrimination further into focus. A case exemplifying the risk attracted a great deal of negative publicity in 2018. The case involved the use of AI software by Amazon to sort job applications, with the lexical analysis of CVs favouring words more commonly used by male applicants, thus discriminating against women (Reuters, 2018).

Concerns around biases and discrimination in AI systems were also apparent in the views of the experts consulted in the Delphi study. The fact that discrimination may be unintentional does not alleviate the ethical concerns, as both intended and unintended algorithmic discrimination can cause harm to those affected. The right to non-discrimination was rated as having relevance exclusively in relation to AI use.

In spite of the ethical concerns, the experts also highlighted the potential of AI algorithms to reduce human biases. AI algorithms can be put to good use when assessing the fairness of human decisions (for example, in relation to recruitment, promotion and pay increases), detecting unconscious human biases and thus contributing to eradicating disparities and inequalities in the workplace.

A need for new digital rights

A wide range of fundamental human rights establish minimum standards for living and working in the EU. However, in recent years, efforts have been made to reinforce some rights challenged by ever more powerful technologies and offer greater protection to citizens (and workers in digital workplaces). For example, through the GDPR, the European Commission sought to create new safeguards to preserve digital privacy and enhance data protection (also enshrined in Article 8 CFREU). Besides the extensive list of data subject rights, the GDPR also introduced the right not to be subjected to automated decisions (Article 22). The much debated right to explanation may be added in future legislation. However, opening the algorithmic black box to ensure the right to explanation is a technically challenging endeavour (requiring non-trivial engineering efforts) and is likely to be opposed by data controllers reluctant to share details of their algorithms to avoid disclosing trade secrets or allowing data subjects to game or manipulate their decision-making system.

Building on the current legal framework, a 2022 European Trade Union Confederation resolution called for an EU directive on algorithmic systems at work, to extend the applicability of key data protection principles to the use of algorithmic systems in the context of employment, and 'to strengthen collective bargaining rights of trade unions as well as information, consultation and participation rights of workers' representatives' (ETUC, 2022).

The right to disconnect is another example of a new digital right, and, at the time of writing in 2023, had been introduced in national legislation in seven EU countries (Belgium, France, Greece, Italy, Portugal, Slovakia and Spain) with a view to counteracting the adverse effects for workers of an always-on work culture enabled by digital technologies (Eurofound, 2023). The right to disconnect is also central to the issue of employee monitoring, as it can help to address the intrusive nature of surveillance technologies, particularly in the context of remote and hybrid working.

The right to disconnect complements the existing fundamental right to fair and just working conditions (Article 31(2) CFREU) and relevant rights provided in secondary law. A 2021 European Parliament resolution considers the right to disconnect 'a fundamental right which is an inseparable part of the new working patterns in the new digital era' and calls on the European Commission to propose a directive 'that enables those who work digitally to disconnect outside their working hours' (European Parliament, 2021).

In the realm of health and safety, the relevant European regulatory framework falls short of addressing new and emerging risks posed by AI and automation technologies. These regulatory gaps are, however, being discussed at EU level. The European Commission has proposed several – mainly non-binding – measures to address the risks that arise from the digitalisation of workplaces in the new 2021–2027 occupational safety and health framework (European Commission, 2021c). On the regulatory front, the Commission has proposed the revision of the Machinery Directive (see European Commission, 2021d) and announced 'further relevant legislative updates' concerning the Display Screen Equipment Directive (Council Directive 90/270/EEC).

In addition, the amended International Labour Organization Declaration on Fundamental Principles and Rights at Work acknowledges that technological innovation may give rise to new occupational risks (in both the physical and psychosocial spheres) and create new challenges for health and safety regulations and their application (ILO, 2022).

Recognising the risks for mental health in increasingly digital workplaces, the European Parliament made more ambitious demands in its resolution on mental health in the digital world of work, calling for EU and Member State lawmakers

> *to strongly commit to actions regulating and implementing a world of digital work which helps to prevent mental health problems, protect mental health and a healthy work–life balance, and reinforce social protection rights in the workplace.*
>
> (European Parliament, 2022a)

Working conditions

From an ethical perspective, the areas of working conditions identified by the experts as most affected by automation and digitisation technologies are related to intrinsic aspects of job quality – particularly skills and autonomy – and health and safety (new physical risks are associated with the use of advanced robots, and new psychosocial risks are linked to AI and IoT use).

AI, IoT and advanced robotics technologies are expected to affect work content and/or methods to varying degrees. Advanced robotics was generally deemed by the experts consulted to improve work organisation by relieving workers from repetitive, physically demanding and menial tasks. There are, however, lingering ethical concerns that advanced robotic applications taking over more complex or challenging tasks may make workers feel deprived of the opportunity to use their acquired skills (or contribute to rapid skills obsolescence) and become a threat to meaningful work.

In the case of IoT use, the experts agreed that it can have both positive and negative effects on work methods and content, affecting aspects of the intrinsic quality of work – for example, levels of autonomy and skills.

While the IoT was deemed by the experts consulted to enable more accurate planning and more efficient workflows, it may also reduce autonomy for some workers. Particularly in production environments, the role of some workers may be reconfigured, requiring them to follow instructions from digitised systems. Workers may also feel that they are under constant supervision due to the extensive data collection and processing capabilities of IoT-based systems. In line with findings from previous Eurofound research (Eurofound, 2021a), a common expectation is that IoT will enable a shift in workers' activities from operational to supervisory, coordination and planning tasks, and from manual to data-driven and more analytical tasks. In their responses, the experts also pointed to the pivotal role of training and reskilling programmes for both blue- and white-collar workers in facilitating the transition to new tasks and job profiles that require more analytical, technical or digital capabilities. Based on the experts' input, there is an ethical dimension to the issue of skills associated with technological change; a lack of adequate training provision makes workers more vulnerable to skills obsolescence and job displacement due to automation and digitisation.

With regard to AI, consensus was reached as to the potential of AI-powered technologies to change the allocation of tasks, which may reduce the number of actions workers can take autonomously. Some experts, however, noted that the use of AI algorithms to support decision-making in the workplace – for example, about workers' task allocation and scheduling – does not necessarily reduce workers' autonomy, as these decisions are typically made by management. From this perspective, the use of AI would, rather, impact the decision-making process (notably, how decisions are taken). Where AI technologies are used to perform management functions, the role of line managers and supervisors may also undergo a profound transformation, with their decision-making powers potentially affected.

While new technologies are typically introduced to make work safer and less demanding, some undesirable or unintended effects may occur after implementation. A change in work intensity is an example of such unintended effects. There was consensus among the

experts that the use of AI, IoT and advanced robots may intensify the pace of work. This largely stems from the technologies' extensive data collection and processing capabilities enabling the quantification of working life (or the datafication of work), prioritising productivity metrics and targets. However, no consensus was reached as to the positive or negative effects of the technologies on working time arrangements and the duration of working time.

In the broad area of health and safety, no consensus was reached in relation to the impact of IoT on the physical work environment. There was nonetheless agreement that the robotisation of the workplace gives rise to new safety concerns due to the increasing complexity of new generations of advanced robots and the greater level of interaction with humans that they require. There is also a risk of advanced robots being hacked or sabotaged and posing risks to workers' safety.

Workers' protection from psychological harm should also be guaranteed in a digitised and automated workplace. In the context of advanced robotics, there was consensus among experts that one risk of increased robotisation is that humans take on a secondary role and are limited to tasks such as supervising robots or machines or exception handling. This may result in workers experiencing a loss of control and a sense of disempowerment and alienation at work.

In situations where the use of digital technologies is extended to more intrusive purposes than initially intended, they can intensify workers' exposure to psychosocial stressors, potentially damaging their well-being and mental health. The experts consulted agreed that both IoT and AI technologies can power more advanced surveillance and monitoring practices, which can have negative psychological effects on workers.

The experts consulted often emphasised that AI, IoT and advanced robotics technologies have, to varying degrees, both negative and positive implications, depending on how the technology is implemented, its purpose and the specific work processes it is applied to, and the type of work concerned.

Ethical principles

Artificial intelligence systems and tools

In the first Delphi study, consensus was reached among the experts consulted on the importance of key ethical principles in relation to the use of AI solutions and systems (including AI-powered advanced robots) in the workplace (see Table 5). These principles, which are consistently found in AI ethics guidelines,[6] are explainability, trustworthiness, transparency, human supervision, and accountability and responsibility. Some of these principles are interrelated, for example transparency and explainability, or cover a range of related concepts, such as the all-encompassing principle of trustworthiness. Consensus over the statements in relation to the principles of explainability, trustworthiness and transparency was reached in the first round of the survey.

Table 5: Ethical principles relevant to the use of AI and proposed statements (on which expert consensus was reached)

Ethical principle	Statement
Explainability	Workers have the right to ask for the parameters included in a decision-making process undertaken by an autonomous machine.
Trustworthiness	In order to ensure the trustworthy use of AI, workers and employers should be protected from potential negative implications of the technologies. To this end, AI technologies should be lawful (respecting all applicable laws and regulations); ethical (respecting ethical values and principles); and robust (both from a technical perspective and taking into account the social environment). More specifically, AI technologies should be designed, deployed and used in such a way that ensures the avoidance of any physical or digital risks that could harm workers.
Transparency	The level of automation introduced by AI technologies and the absence of transparency concerning their functioning could interfere with workers' freedom of self-determination and their ability to truly consent.
Human supervision	AI processes and decisions should be subject to human supervision. Supervisory control functions should enable workers to have a thorough understanding of AI technologies' operations, with a view to enabling effective conformity assessments of AI-based software and machines incorporating it.
Accountability and responsibility	Via the integration of AI tools and applications, some machines have new responsibilities and even autonomy, as well as being expected to display various forms of human intelligence and make decisions about workers themselves. In addition, automated decision-making technologies can lead to unwanted or erroneous decisions, which introduces questions about who is legally responsible for any damage or any other undesirable outcome that emerges.

Source: *First Delphi study, 2021*

6 An extensive review of AI ethics guidelines by Jobin et al (2019) identified the following ethical principles: transparency, justice and fairness, non-maleficence, responsibility, privacy, beneficence, freedom and autonomy, trust, sustainability, dignity and solidarity.

With regard to the principle of accountability and responsibility in relation to AI use, most experts took issue with the second part of the proposed statement on liabilities in the event of unwanted or erroneous decisions. While the experts considered that machines should not be held accountable, the question remains as to who – the employer, the producer of the AI tool or application, or even the employee who uses the tool – should be held accountable. Several features of digital technologies and particularly AI systems – such as their high levels of complexity and connectivity, their autonomous or semi-autonomous capabilities, the opacity of their internal workings and their unpredictability – pose challenges to the application of liability rules (Expert Group on Liability and New Technologies, 2019; European Commission, 2021e).

Although consensus was reached over the principle of trustworthiness, the views of the experts somewhat diverged from the proposed statement (which was largely based on the European Commission's *Ethics guidelines for trustworthy AI)*. While there was agreement that AI-based systems and tools should be lawful and robust, some experts contended that such systems are not per se ethical or unethical. What makes them ethical is the way humans use them and the context of their application.

A relevant ethical principle identified by the experts in the first Delphi study concerned human supervision. Part of the statement was, however, contested by most experts. They argued that having 'supervisory control functions' does not necessarily enable workers to have a thorough understanding of the internal workings of AI technologies. For a variety of reasons, a thorough explanation might not be possible; computations could perpetually change based on the data gathered; and, in some cases, humans may have difficulties keeping up with the speed of AI systems.

Focusing on machine learning applications

In a follow-up Delphi study with technology experts, which focused more specifically on machine learning applications, the experts consulted agreed on a broader range of relevant ethical principles (Table 6). These were non-discrimination, fairness, prevention of harm, trustworthiness, transparency, human in the loop (human supervision), explainability, accountability and responsibility, human autonomy and self-determination, and data semantic integrity.

Many of these principles are interconnected; for example, transparency can be regarded as a requirement for explainability, accountability and responsibility, and is also linked to fairness. It is essential to understand how a system works to be able to explain to those affected by a decision how it came to that decision and establish who is accountable for it, especially if the technology goes awry or is misused. The lack of mechanisms to determine accountability makes it difficult for regulators and courts to determine liabilities where technologies violate human rights and ethical principles. While in the context of AI accountability refers to the determination of (legal and social) liabilities for decisions made, the concept of responsibility may be regarded as more 'internally motivated' (Vakkuri et al, 2020).

Consensus was reached on most of the proposed statements in the first round of the Delphi survey. The principles of human autonomy and self-determination, non-discrimination and data semantic integrity were proposed in the second survey round and consensus was reached.

Table 6: Ethical principles relevant to the design and use of machine learning applications and proposed statements (on which expert consensus was reached)

Ethical principle	Statement
Non-discrimination	Programmed machines can have inherent biases and can classify humans into artificially restrictive groups or identities.
Fairness	The development, deployment and use of AI systems must be fair. Fairness has both a substantive and a procedural dimension. The substantive dimension implies a commitment to ensuring the equal and just distribution of both benefits and costs and ensuring that individuals and groups are free from unfair bias, discrimination and stigmatisation. The procedural dimension of fairness entails the ability to contest and seek effective redress against decisions made by AI systems and by the humans operating them.
Prevention of harm	AI systems should neither cause nor exacerbate harm or otherwise adversely affect human beings. This entails the protection of human dignity as well as mental and physical integrity. AI systems and the environments in which they operate must be safe and secure. Preventing harm also entails the consideration of the natural environment and all living beings.
Trustworthiness	Trust in the development, deployment and use of AI systems concerns not only the technology's inherent properties, but also the qualities of the sociotechnical systems involving AI applications. Trustworthy AI has three components, which should be met throughout the system's entire lifecycle: lawful, ethical and robust.
Transparency	Processes need to be transparent, the capabilities and purpose of AI systems openly communicated, and decisions – to the extent possible – explained to those directly and indirectly affected.

Ethical principle	Statement
Human in the loop (human supervision)	Human oversight helps ensure that an AI system does not undermine human autonomy or cause other adverse effects. The 'human in control' principle requires that humans, and not algorithms, should ultimately be in control of, and therefore responsible for, relevant decisions.
Explainability	Explainability concerns the ability to explain both the technical processes of an AI system and the related human decisions – for example, the areas of application of a system. Technical explainability requires that the decisions made by an AI system can be understood and traced by human beings.
Accountability and responsibility	The principles of accountability and responsibility are closely linked to the principle of fairness. Mechanisms must be put in place to ensure that responsibility and accountability for AI systems and their outcomes can be attributed, both before and after their development, deployment and use.
Human autonomy and self-determination	Workers may lose autonomy, as algorithms can take decisions instead of them.
Data semantic integrity*	Data collected in one context must not be used to train systems for use in a different context where the observations and predictions are operationalised differently.

Note: * *Data semantic integrity is related to the data protection principle of purpose limitation (Article 5 GDPR).*
Source: *Second Delphi study, 2022*

No consensus was reached in relation to the principle of 'privacy by design' or 'privacy by default'. However, this was not because the experts deemed it unimportant. Rather, it was a matter of adding nuance to the proposed statement that they were asked to agree on, which overemphasised compliance with legal requirements. Overall, the principle of privacy by design is regarded as a promising way to promote privacy and ensure that developers pay attention to its embeddedness in technology design rather than viewing it merely as a desirable outcome or a compliance issue. Some experts noted that privacy by design should become an integral part of technology design and use, underpinned by compliance with meaningful regulations and a corporate culture with an emphasis on ethical values; otherwise, it will simply serve as a slogan or to facilitate ethics washing.

Internet of things

In the first Delphi study, the participants were also consulted on ethical principles at stake in relation to the use of the IoT in the workplace. Some of the ethical principles flagged as relevant to AI use – namely transparency, and accountability and responsibility – were deemed equally important in the context of IoT use, particularly if the technology is used to power surveillance or monitoring systems. Consensus was reached on the relevance of the principle of transparency in the first round of the survey, whereas experts identified another two ethical principles – accountability and responsibility, and trust – in the first round and reached consensus on them in the second round (Table 7).

The principle of transparency relates to the need to make workers aware of the purposes of processing collected data and provide them with meaningful information. With regard to the principle of trust, the experts consulted considered that informing workers about the use of IoT technology is necessary to ensure their trust in the technology, but is not sufficient on its own. From the workers' perspective, the extent to which the technology can be trusted depends on how the information gleaned from the IoT is fed back to workers and whether it is done without questioning their competence, commitment and honesty.

Table 7: Ethical principles relevant to the use of IoT and proposed statements (on which expert consensus was reached)

Ethical principle	Statement
Transparency	The pervasiveness of IoT sensors and the invisibility of the control and supervision enabled by IoT-based monitoring systems makes it more difficult for employees to contest management decisions based on sensor-collected data.
Accountability and responsibility	Automated decision-making technologies powered by data gathered via IoT sensors can lead to unwanted or erroneous decisions, which introduces questions about who is legally responsible for any damage or any other undesirable outcome that emerges.
Trust	For the use of IoT technologies to be trustworthy, workers should be informed about their use in the workplace and how the technologies are used to their benefit. The use of IoT for monitoring and surveillance practices might lead to mistrust among workers vis-à-vis their employer or management.

Source: *First Delphi study, 2021*

The lack of consensus among the experts over the relevance of other ethical principles to IoT use in the workplace tended to relate to general disagreement over the proposed definitions or statements underpinning such principles. One notable example concerned the principles of fairness and justice. These are distinct but related principles often discussed by ethicists. They are also related to the fundamental right to equality and non-discrimination (Article 21 CFREU) and the right to access to an effective remedy and to a fair trial (Article 47 CFREU).

Ethics applied to technology design, development and use

General considerations

There are intrinsic difficulties in the straightforward application of ethical principles to the development and use of new and emerging technologies.

Due to the novelty and versatility of digital technologies, they may raise new (and unprecedented) ethical concerns that cannot be fully addressed within pre-existing frameworks or moral and legal rules. For example, the increasing autonomous capabilities of robotic applications in digitally connected work environments may require current rules and legal provisions to be revisited, with a view to clarifying accountability and responsibilities at different levels and, not least, safeguarding workers' health and safety.

An additional difficulty is that the effects of emerging technologies – which, by definition, are not fully mature – are not (yet) visible. As new technologies carry a certain degree of unpredictability, some scholars advocate for a precautionary approach to their regulation to limit or prevent undesirable risks and unintended effects (Wolff, 2014; Kendal, 2022).

An ethical dilemma may also arise as a result of the context in and purpose for which a new technology is used, rather than the technology itself. For example, digital tracking technologies may be ethically desirable when used to provide greater security for workers operating in hazardous environments. By contrast, the use of such technologies for the sole purpose of monitoring workers' presence and tracking their productivity is ethically questionable.

Another challenge is related to situations where there is conflict between ethical principles and there is no guidance on how to deal with such conflict (as in most ethical guidelines, principles and values are non-hierarchical). For example, the use of digital technologies in many workplaces during the pandemic, ostensibly to guarantee the safety of workers, has raised some ethical concerns in relation to potential breaches of workers' data protection and privacy rights (Eurofound, 2022a). In such situations, one viable approach is 'to determine which rights, principles, and duties carry the most ethical force and/or find a solution that minimises the ethical violation' (Vallor et al, 2020).

Furthermore, the standard approach to ethics, generally taken in biomedical ethics,[7] presupposes an ethically laudable goal – that is, to contribute to human welfare. This logic may not be entirely transferable to the ethics of technology, as technology may not necessarily be developed in pursuit of an intrinsic ethical goal (Stahl, 2021). This suggests that ethical questioning and an ethics assessment may be required at the research and development and design stages of technology development to avoid some ethical pitfalls. One such example is the 'value-sensitive design' approach, which integrates human values and principles in technological design and development (Friedman et al, 2008; Manders-Huits and van den Hoven, 2009).

Embeddedness of ethical principles in technology design and development

Based on the experts' views, most of the principles identified as having ethical relevance in the context of machine learning applications can be embedded (at least to some extent) in technology design and development. This embeddedness, however, poses some difficulties due to the inherent characteristics of machine learning. AI systems are complex; they are based on intricate sequences of feedback loops, which make it difficult to predict the outcomes, even when all parameters are known. This calls into question the complete applicability of several ethical principles, particularly the extent to which humans can be truly in the loop, and whether the transparency, explainability and more broadly trustworthiness of such systems can be fully ensured. Business ethicists, however, question the common assumption that machine learning is inscrutable, as it may serve as a justification for transferring responsibility for wrongful or harmful decisions to autonomous AI agents; in this regard, D'Cruz et al (2022) argue that 'digital technologies designed to be inscrutable are more about corporate power than any design requirement'.

For other scholars, the transparency requirements vary depending on the needs of the stakeholders involved (Weller, 2019). For example, while access to the source code may be sufficient for a developer to verify that the system works as intended, this would not satisfy the

7 In biomedical ethics, practices and rules can be justified by the application of four key principles: autonomy, non-maleficence, beneficence and justice.

requirements for transparency and for the technology to be explained to workers or their representatives. Looking into the issue of the transparency of machine learning, Buiten (2019) suggests that 'regulatory transparency requirements should be context-dependent and based on risks to safety, fairness, and privacy'.

No consensus was reached in the Delphi study over the embeddedness of the non-discrimination principle in the design of machine learning systems. While recognising that biases could be partly fixed through data minimisation and using statistical techniques to assess and correct biases in data, most experts contested the idea that algorithmic discrimination remains a complex issue that must be addressed solely in technology design and development. This is because algorithmic biases mirror structural inequalities in society, which is reflected in the training data feeding AI systems. These data can be unrepresentative, incomplete or flawed. Another important consideration is that algorithms also reflect the perspectives, normative values and biases of the developers or data scientists responsible for their development, no matter how well intentioned they are or how objective they strive to be. Furthermore, those who design and develop algorithms are affiliated to software companies with vested interests, not necessarily aligned with the interests of workers whom algorithmic decisions affect, or are simply not concerned with developing technologies for social good. In the face of these challenges, both internal control mechanisms and external audits performed by expert independent bodies are required to deal more effectively with the risk of discrimination – as well as other risks – related to the use of machine learning models in both the public and the private sectors.

The technology experts consulted highlighted several pre-conditions and requirements for embedding the identified ethical principles in technology design and development, suggesting that there is no silver bullet for ethical AI design and use. The pre-conditions and requirements selected most by the experts included the training and education of developers and managers on AI ethics to cultivate their ethical sensitivity; the involvement of multidisciplinary teams in AI design and development; the creation of ethical criteria and procedures to regularly check compliance with these criteria; and the establishment of a governance framework including mechanisms for redressing harm caused by algorithmic activity.

A range of instruments and tools were identified by the experts consulted to identify ethical pitfalls and mitigate the potential negative implications of AI technologies. These include, for instance, algorithmic bug bounties, whereby participants are rewarded for identifying exploitable flaws or bugs in software, and the measurement of biases in algorithmic predictions and corrections using pre-processing, in-processing or post-processing techniques. A range of tools are available for auditing the quality of datasets used in algorithmic decision-making systems, as documented in a 2022 report by the European Parliamentary Research Service. AI governance models and auditing tools are also being developed as part of both public and private initiatives. Examples include the Oxford Internet Institute's programme developing a meta-toolkit for trustworthy and accountable AI, the Artificial Intelligence Governance and Auditing programme by a University of Turku consortium building AI governance mechanisms and models, and the AI Fairness 360 toolkit developed by IBM.

Other useful tools for assessing the ethics of technologies are ethical matrixes, applying a set of relevant ethical principles to each affected group, and ethical dilemma scenarios, prompting ethical thinking among different stakeholders at the technology design stage (Wright et al, 2014). Ethics canvases, ethical assessments and ethical risk analyses (and risk–benefit analyses) have also recently emerged as tools to identify the ethical risks of emerging technologies and are used to inform risk management strategies. Furthermore, the United Nations Guiding Principles on Business and Human Rights are also deemed relevant to the design, development and use of technology in the workplace. Guiding principle 18 states that a human rights risk assessment should be performed as a first step in gauging the risks associated with business activities; the results of the assessment should feed into the design of risk prevention and mitigation measures (United Nations, 2011).

Morley et al (2020) have proposed a typology for assessing available ethics tools and methods in terms of high-level ethical principles (beneficence, non-maleficence, autonomy, justice and explicability) and the stages of development of machine learning systems in which they are used. The authors point to the lack of usability of most of the identified tools and methods, which disincentivises developers' use of them, and their uneven distribution across both development stages and identified ethical principles (Morely et al, 2020).

AI ethics has overall received limited attention in the literature on engineering, with most research in the field being conceptual and theoretical. Despite the proliferation of AI ethics guidelines,[8] AI ethics is not

8 Among the most notable examples of guidelines assisting in the ethical design of intelligent systems are the Institute of Electrical and Electronics Engineers guidelines for ethically aligned design, developed as part of the institute's Global Initiative on Ethics of Autonomous and Intelligent Systems.

systematically considered in software engineering projects (Vakkuri et al, 2020). When it comes to identifying ethical pitfalls of the applications being designed and developed, the experts consulted indicated that this is typically done by developers and data scientists who know intuitively what issues present ethical concerns and are aware of policies and ethical guidelines, rather than through systematically ethical risk assessments.

The experts pointed out the benefits of mainstreaming ethics in training and standard courses on computing and engineering. This will contribute to fostering ethical sensibility and practice and creating awareness of ethical concerns in the early stages of technology design and development.

The experts consulted ultimately called for ethical principles to be embedded in the whole lifecycle of AI systems, and not only at the design and development stage. The purpose for which a technology is used once developed is also relevant. A technology can be designed for one ethical purpose and then reused for another ethically questionable purpose. Therefore, an AI governance framework must be established for the operational management of the technology as well as tight control and supervision measures, to ensure that ethical principles and standards are upheld throughout the process from design and development to implementation and use in the workplace.

Key takeaways

- Although new digital technologies present many benefits and opportunities, they also bring to the fore new ethical issues, reshaping the meaning and understanding of ethics in the digital workplace. According to the experts consulted, digital technologies for work have many ethical implications, with ramifications extending to several areas of working conditions and fundamental rights and involving an ever-expanding set of ethical principles. Based on the experts' views, impacts on job quality, human rights and ethical principles should be considered as part of a comprehensive assessment – not limited to GDPR compliance – and as a first step in the process of implementing digital technologies, to ensure the effective and forward-looking prevention and management of ethical risks.
- With digital technologies increasingly taking a central role in working life, the fundamental rights framework plays an important part in the regulation and application of technologies in the workplace. Technological change also raises the questions of whether the current framework provides an appropriate basis for responding to the new ethical and legal challenges posed by digital technologies and, if so, how effectively fundamental rights can be enforced and safeguarded in a technology-driven work environment.
- According to the Delphi study conducted with technology experts, most of the ethical principles consistently mentioned in ethics guidelines can be, at least to some extent, embedded in the design and development of machine learning applications. Whether technology solutions are adopted by private or public organisations, ethical issues should be considered throughout the whole process, from the design and development of technologies to their deployment and use in the workplace. If a technology is deployed for a different purpose or in a different context from that for which it was initially designed and used, new ethical issues may surface. These need to be carefully assessed and dealt with. This entails the adoption of an iterative framework whereby ethical considerations are embedded throughout the whole technology lifecycle to make sure that technologies are designed, developed and used for good purposes.
- Notwithstanding the importance of reaching consensus over the meaning of ethical principles in the context of workplace digitalisation, there is a need to bridge the existing gap between theory and practice and move the discussion on ethics beyond the definitions underpinning ethical principles. To this end, an important step is fostering a broad dialogue on the ethical digitalisation of work involving experts from different disciplines, policymakers, social partners and civil society organisations, with a view to coordinating efforts to develop tools and practical guidance for ethical technology design and use.

2 Debates and policies on digitalisation of work and ethical issues

European Union policy context

In recent years, the EU has made efforts to take the lead in regulating the use of digital technologies in society. Most of the recent regulatory efforts are underpinned by the need to protect fundamental rights in the face of fast technological developments.

Back in 2015, the European Data Protection Supervisor (EDPS) published Opinion 4/2015, entitled *Towards a new digital ethics*. The opinion identified several technological trends, such as big data, the IoT and networked sensors, and AI. According to the EDPS, the use of these technologies – if left unchecked – poses unprecedented challenges to the principles of data protection and more broadly human rights. In its opinion, the EDPS called for an updated legal framework underpinned by ethical considerations. The GDPR was a step in this direction and it is to date recognised globally as the blueprint for data protection and digital privacy. Many principles of the GDPR are at their core ethical principles.

Since the GDPR entered into force in 2018, several European initiatives on ethical AI have emerged, laying the groundwork for the European Commission's proposal for a regulation on a European approach to AI (known as the AI Act). The proposed AI Act builds on the work of the AI HLEG, appointed in 2018 by the European Commission. The AI HLEG put forward the concept of trustworthiness in relation to AI systems, which is anchored to three characteristics: being lawful, ethical and robust. An important milestone in the work of the AI HLEG was the publishing of its *Ethical guidelines for trustworthy AI*, outlining seven core requirements that AI systems should meet to be determined as trustworthy. These are human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental well-being; and accountability.

The EU's strong commitment to promoting shared European values and ensuring respect for human rights in AI development and application was reaffirmed in the European Commission's white paper on AI.

> *The Commission is committed to enabling scientific breakthrough, to preserving the EU's technological leadership and to ensuring that new technologies are at the service of all Europeans – improving their lives while respecting their rights.*
>
> (European Commission, 2020a, p. 1)

In the same year, the European Parliament adopted a resolution with recommendations to the Commission on a framework of ethical aspects of AI, robotics and related technologies, calling for 'an effective, comprehensive and future-proof regulatory framework' (applicable to emerging technologies) for the safeguarding of fundamental human rights and ethical values.

The sustained public debate around the need to consider legal and ethical aspects in technology use – particularly AI use – culminated with the publication in April 2021 of the draft AI Act. The proposed regulation is intended to regulate the systems that use AI (not the technology itself) through a risk-based approach grounded in transparency, trustworthiness and explainability.

The draft AI Act differentiates between unacceptable risk, high risk and low or minimal risk. While banning outright the use of AI systems that pose unacceptable risks,[9] the regulation mainly focuses on high-risk applications deemed to impinge on fundamental rights. In the context of employment, high-risk AI systems include those used for work management and recruitment. Such systems must undergo a conformity test (before reaching the single market) and be compliant with prescribed risk management system, transparency, human oversight, data quality and cybersecurity requirements. Furthermore, the draft compromise text agreed by the European Parliament's Committee on Internal Market and Consumer Protection and Committee on Civil Liberties, Justice and Home Affairs also proposes that worker representatives are consulted prior to the adoption of a high-risk AI system in the workplace, 'with a view to reaching an agreement' (European Parliament, 2023).

9 In 2021, the European Data Protection Supervisor (EDPS) and the European Data Protection Board (EDPB) published Joint Opinion 5/2021 on the AI Act, calling for an outright ban on 'any use of AI for automated recognition of human features in publicly accessible places … in any context' (EDPB and EDPS, 2021, p. 11). Such practices are deemed not to meet the necessity and proportionality requirements for the processing of personal data.

At the time of writing, the EU and the US have also announced that a voluntary (interim) code of conduct on AI will be created to provide some safeguards and protection until the new legislation comes into effect, thus bridging a policy gap that widens as technology develops at a faster rate than expected (Reuters, 2023).

To complement such efforts, the European Commission has also proposed an AI liability directive, which sets out the rules for accessing redress in instances of harm caused by AI systems (European Commission, 2022). There is evidence that liabilities in relation to AI use are indeed a key concern for enterprises considering adopting AI technologies. In an EU-wide enterprise survey conducted in 2020,[10] 33% of the surveyed enterprises indicated 'liability for potential damages' as a major barrier to AI use (European Commission, 2020b).

The upcoming AI Act (and complementary legislation) is likely to have important effects on global markets (Brookings, 2022). It is, however, yet to be seen whether the new legislation will generate the so-called Brussels effect, as the GDPR did, becoming a global benchmark for regulating the design and application of AI systems.

However, trade union organisations, civil society organisations and labour law scholars have pointed to some gaps and weaknesses in the proposed provisions on high-risk uses of AI, including in the context of employment (Ponce del Castillo, 2020; De Stefano and Taes, 2023; Future of Life Institute, undated). The risk classification of AI systems is, for example, regarded as problematic, as applications may generate risks depending on the specific context and purpose for which they are used. Some scholars also warn that, if adopted, the AI Act may enable deregulation and prevail over more restrictive national legislation (De Stefano and Taes, 2023).

Topics in national policy debates

Digital technologies are generally discussed positively and from a broad perspective in the national policy discourse. Policy stakeholders recognise the significant benefits that digitalisation can bring to the economy and society at large. Such benefits include, for example, increased productivity and competitiveness, improved public service delivery, a faster transition to net zero emissions and, overall, better functioning societies.

Both the review of policy documents and the interviews with policy stakeholders conducted as part of this study nonetheless suggest a growing awareness among policymakers about some negative implications of digital technologies, particularly in the context of AI use.

The increasing attention on potential ethical pitfalls of AI has been partly fuelled by the ongoing discussions at European level on AI regulation as well as the increasing use of AI for ethically sensitive applications (for example, for HR and work management or public service delivery and recent developments related to ChatGDP).

However, when it comes to employment and work, the range of ethical concerns raised in policy debates are limited to data protection and privacy, transparency, discrimination and potential biases.

In relation to data protection, the stakeholders consulted often emphasised the need for transparency with respect to who owns and processes the data, the need to ensure the anonymity of data used by AI algorithms, and the importance of defining what type of data can be collected and for what purposes. The increase in attention on and awareness of the importance of ethical principles underpinning data protection requirements in national policy debates is largely due to knowledge acquired in the context of the national implementation of the GDPR, and greater familiarity with EU-wide standards for the use and processing of personal data and potential ethical pitfalls of non-compliance.

Issues around discrimination and biases in the context of AI use are usually discussed in relation to requirements for greater transparency and accountability in AI-informed decision-making. Stakeholders consulted emphasised the potential issues raised by AI technologies due to the use of biased historical datasets for training algorithms and the use of novel types of data (for example, biometric data and social media data) that may lead to unfair decision-making in the context of employment. The stakeholders therefore called for policies ensuring greater transparency in AI use and addressing the sources of bias to prevent AI systems reproducing or exacerbating existing inequalities in both society and the workplace.

Other ethical principles routinely mentioned by ethics experts that should inform the regulation of AI design and use surface less frequently in policy debates. Furthermore, issues around responsibility and accountability in AI use, and human supervision, tend to be discussed in broad terms and often in relation to actions aimed at increasing public trust and acceptance of AI. Examples of this type of action include publicly available general courses, such as the Finnish 'Elements of AI' course, that introduce AI technologies and their societal implications to a broad audience.

10 The survey was conducted in 2020 on behalf of the European Commission by Ipsos/iCite in the EU27, Norway, Iceland and the UK using Computer Assisted Telephone Interviewing (CATI) and reaching a total of 9,640 enterprises across the 30 countries surveyed. The target respondent in each enterprise was broadly defined as an employee who knows about the use of technology in the company.

In terms of the implications of digitalisation for working conditions, adaptation to new skills requirements is the most frequently discussed topic in national debates. Concerns in this area have intensified due to the increasing labour shortages in sectors vital to the digital transition, such as the information and communications technology (ICT) sector, and expectations that the further deployment of AI technologies will have a major impact on the job market by changing occupations and skills needs (Eurofound, 2023). It is predicted that these changes will generate future labour shortages.

There was broad consensus among the stakeholders consulted as to the need to invest further in training provision and competency development for workers on a lifelong basis, with a view to increasing workers' resilience in a fast-changing, technology-driven labour market.

Governments', employers' and trade unions' perspectives

Governments

As evidenced by strategic policy documents issued by public bodies, digital technologies are generally regarded as a source of innovation for economic growth and sustainable development. In most of the selected countries, public investments are being directed to boost innovation and improve digital infrastructure.

The technological transformation of workplaces was consistently viewed positively by the government representatives consulted. They saw it as an opportunity to optimise and reorganise work processes, increase workplace safety and complement or augment workers' capacities. However, while recognising the benefits of workplace digitalisation, government representatives drew attention to pressing challenges requiring strong policy responses. One such challenge is that technological change affecting work methods and content and redefining tasks will continue to drive changes in skills requirements. These changes can be detrimental to workers if not paired with the adequate provision of training, upskilling and reskilling programmes.

Several government representatives also highlighted the importance of building trust in AI systems as a cornerstone of the digital transformation of society. The issue of trust is generally associated in the political discourse with considerations around transparency, accountability and fairness or non-discrimination. The 2020 District Court of the Hague ruling against the Dutch government's use of an AI risk profiling system called System Risk Indication (SyRI) for social welfare fraud detection brought to light major ethical pitfalls of AI systems with regard to transparency and accountability. In the judgment, the Dutch court relied on the principles of transparency, purpose limitation and data minimisation set out in the GDPR to establish a violation of Article 8 (on the right to privacy) of the European Convention on Human Rights (Rachovitsa and Johann, 2022).

There have been other cases of public organisations being incautious in their use of AI to provide public services. For example, in Finland, the national public broadcasting company Yle faced media backlash for using an AI moderation system that did not weed out offensive comments from online discussions, and they were subsequently compelled to revert to human moderators (Yle, 2020).

The increasing use of automated algorithmic systems in municipalities in Sweden to deliver social services has also sparked an intense debate on the transparency requirements of such systems. The Swedish union for academics in social sciences (Akademikerförbundet SSR) reported the municipality of Trelleborg to the Ombudsperson for its lack of transparency in the use of an automated welfare distribution system and noted that 'it becomes increasingly clear that trust in the public sector is dependent on transparency and comprehensibility' (Sydsvenskan, 2020).

Several governments have initiated discussions and reflected on policy responses appropriate for dealing with the new risks posed by AI. For example, they have discussed whether existing legislation needs to be complemented or reviewed and new mechanisms put in place to ensure the conformity of algorithmic systems with ethical and legal standards.

Data protection continues to feature high on government agendas, particularly in Germany and France. According to the government representatives consulted, the prevailing debate is whether existing regulations, ethical codes and soft law provide an adequate framework for ensuring data safety and dealing with issues around biases in machine learning algorithms. In this regard, Finnish government representatives emphasised the difficulty of developing universally applicable ethical guidelines for AI systems, suggesting that a more viable approach would be to develop ethical guidelines on a case-by-case basis. These would consider the specific features of each technology, with ethical reviews or assessments accompanying the process of technological development (Finnish Government, 2018).

In several countries, the unprecedented surge in remote working during the COVID-19 pandemic brought to the attention of most governments the need to modernise legislation on teleworking and strengthen protection and safeguards for workers in digital work environments. According to various government representatives, this is an urgent priority, as remote work is becoming a more established working arrangement than it was in pre-pandemic times.

The series of legislative initiatives recently implemented by the Spanish government demonstrates its commitment to modernising the legal framework to ensure that technological progress does not come at the expense of workers' rights. In this vein, the Spanish Minister for Labour, Yolanda Díaz, has stated that

> *Algorithms are not abstract entities, there are processes behind them that must be analysed and assessed, monitored. The employer's control of labour activity, for example, is a right, but it is not unlimited. And the protection of fundamental rights, data protection and the guarantee of digital rights must always take precedence.*
>
> (El País, 2020)

In addition, German and Swedish government representatives interviewed for this study emphasised that the increase in the uptake of remote working due to the COVID-19 pandemic has heightened ethical concerns about workplace digitalisation, particularly with regard to increased surveillance and monitoring of workers. This concern has many potential ramifications in terms of negative impacts on working conditions (including with effects on mental health) and fundamental rights.

A study by the Swedish Work Environment Authority (2015) highlighted increased surveillance at work as an area of concern, along with other (related) negative effects of workplace digitalisation, such as greater expectations on workers to always be available, blurred work–life boundaries and work intensification. Surveillance and work intensity have also been topics of discussion within the special Enquete Commission on AI, set up in 2018 by the German parliament and tasked with investigating the potential impact of AI on the world of work.

Belgian government representatives have a more neutral stance on the topic. When it comes to teleworking and remote working, their main concern is ensuring a good work–life balance, rather than counteracting the increase in surveillance at work. In this regard, in 2018, Belgium enacted legislation containing provisions on the right to disconnect and requiring employers (in companies with more than 50 employees) to address the right to disconnect in collective agreements or, in the absence of such agreements, in their work regulations (Eurofound, 2021c).

Employer organisations

Employer organisations perceive digitalisation and AI as an important opportunity for business growth and innovation. From their perspective, the use of new technologies in the workplace is expected to generate dividends in terms of both competitiveness and productivity as well as broader societal prosperity. For example, the Confederation of German Employers' Associations (Bundesvereinigung der Deutschen Arbeitgeberverbände, BDA) argues that the economic opportunities brought by AI are significant and that 'technology leadership is a decisive competitive advantage that secures employment and prosperity in the long term and enables standards and values to be set' (BDA, 2020).

To leverage the opportunities provided by AI, employers tend to advocate for two types of actions, one entailing investments in digital infrastructure and the other channelling resources into upgrading the skills of the workforce. Employer organisations interviewed in the selected Member States observed that one key implication of the digitalisation of work is the emergence of new skills requirements, which will continue to generate significant pressure on the labour market. Securing investments in upgrading skills is therefore a key priority on their agenda. Such investments should target the broader educational curricula to better prepare young people for increasingly digital workplaces, as well as tailored programmes geared to respond to existing shortages of ICT professionals.

In Finland, Germany, the Netherlands and Sweden, employer organisations also advocate for regulatory changes to facilitate investments and innovation in the field of AI. For example, the BDA argues that a flexible legal framework that allows companies to adapt to technological change is a prerequisite for the successful uptake of AI. According to the German employer organisations consulted for this study, a key priority for employers is harmonising legislation between different federal states to reduce complexity and legal uncertainty for businesses. From a broader perspective, the Confederation of Swedish Enterprise (Svenskt Näringsliv) argues that 'large and thoroughgoing changes in society will inevitably create calls for legislation … which needs to be based on empirical knowledge, and that can often only be attained in retrospect' (Confederation of Swedish Enterprise, 2018, p. 1).

Although not a core focus, employer organisations in some countries recognise that digitalisation, specifically AI, can create risks related to cybersecurity and privacy as well as a lack of transparency and potential biases. In Denmark, employer representatives interviewed argued that ethical standards must be upheld in instances when workers' data are collected by digital systems and that data collection practices restricting individuals' behaviour should be avoided. In Sweden, the position of the Confederation of Swedish Enterprise on AI is that 'AI applications could eventually cause redundancy in a number of professions' but also 'make it possible to systematically … influence the preferences of individuals in a variety of ways' (Confederation of Swedish Enterprise, 2018, p. 6). While acknowledging that technologies can have both positive and negative effects, some employer

organisations argued that it may be ethical to use certain technologies, and they may even facilitate the fulfilment of human rights, for example in situations where technologies are essential to safeguarding workers' health and safety. The use of certain technologies may be also ethical if they make work easier and less demanding, especially for certain groups – for example, people with disabilities.

While recognising that legislation is essential to avoid technologies being misused, employer organisations warned against overregulation, which may hinder innovation. In Finland, the Netherlands, Spain and Sweden, employers advocate instead for the use of collective bargaining and sectoral supervision to address the potential risks posed by the use of AI. The Spanish Confederation of Employers' Organizations (Confederación Española de Organizaciones Empresariales, CEOE) recommends strengthening social dialogue in response to the impact of digitalisation on work (Godino et al, 2022). In addition, the Confederation of Netherlands Industry and Employers calls for sectoral supervision of AI and continuous dialogue between public authorities and stakeholders, with a focus on providing guidance, and exchanging information and good practices (VNO-NCW, 2022). Notwithstanding the importance of trust and transparency in the use of data processing technologies, the BDA considers that suitable rules and procedures ought to be developed at establishment level and through collective bargaining.

Trade unions

Trade union representatives consulted in this study expressed similar concerns to employer organisations as to the necessity to invest in upgrading the skills and developing the competencies of the workforce. From their perspective, technological unemployment and unequal access to the labour market could be potential outcomes of digitalisation, especially for specific categories of workers (for example, blue-collar workers). There is, however, an understanding that training workers and upgrading their skills – albeit useful – is only part of the solution. Trade unions also advocate for the involvement of staff from an early stage in technology adoption to help build trust among employees, reduce workers' apprehension about digital disruption and create positive attitudes towards change.

The range of concerns arising from workplace digitalisation expressed by the trade union representatives consulted is much wider than those highlighted by other stakeholders. For example, a source of concern relates to the expanded surveillance capabilities of most new digital technologies and the negative implications of these technologies for working conditions and their potential interference with data protection and privacy rights. Workplace surveillance is also regarded as deterring workers from exercising their right to collective representation.

Representatives of Belgian, Spanish and Swedish trade unions also warned against the negative effects of an emerging data-driven work culture – enabled by ubiquitous technologies – whereby workers are assessed based on pre-established performance metrics and are continuously visible to their employers. According to several trade union representatives, such data-driven management practices give rise to new psychosocial risks that harm workers' mental health and can potentially dehumanise work.

The trade union representatives also voiced concerns over discriminatory and biased algorithmic decision-making, and liabilities and responsibilities in the event of unfair outcomes or harm caused by algorithmic activity. Such concerns relate primarily to the use of AI systems in human resource management practices, particularly recruitment and performance appraisals.

The principles of transparency and explainability are often invoked by trade unions when it comes to AI systems. Transparency requirements are essential to establish responsibilities, especially for harm or offence caused by decisions made by these systems. According to the trade union representatives consulted, such requirements should go beyond the mere disclosure of information and rather ensure that algorithms are 'explainable'. The principle of explainability is, however, still debated, and it remains a challenge to explain the internal workings of an algorithm and the logic of how decisions are made due to the inherent complexity and unpredictability of machine learning systems.

Another concern frequently mentioned by trade union representatives relates to data protection and privacy, including the type of employee data collected by AI systems and the purpose for which these data are used. Concerns around the potential misuse of employee data and issues around privacy are very prominent in policy debates, particularly in France, Germany and Poland. The Belgian trade unions consulted also emphasised the negative impact of increased monitoring and surveillance on collective bargaining, arguing that increased surveillance through digital means could contribute to the erosion of collective bargaining rights.

Stakeholders' perspectives across EU Member States

Table 8 outlines the main areas of concern reported by governments, employer organisations and trade unions across the selected EU Member States.

Table 8: Areas of concern reported by policy stakeholders consulted in the selected EU Member States

Areas of concern	Government(s)	Employer organisations	Trade unions
Skills development and adaptation	Belgium, Finland, France, Germany, Netherlands, Poland, Spain, Sweden	Belgium, Denmark, Finland, France, Germany, Netherlands, Poland, Spain, Sweden	Belgium, Denmark, Finland, France, Germany, Netherlands, Poland, Spain, Sweden
Data protection and privacy	France, Germany, Netherlands, Spain, Sweden	Denmark, Poland, Sweden	Belgium, Finland, France, Germany, Poland, Sweden
Transparency, accountability and trust	Belgium, France, Germany, Netherlands, Spain, Sweden	Denmark, Finland, Poland, Spain, Sweden	Denmark, Finland, Poland, Spain, Sweden
Discrimination and potential biases	Netherlands, Sweden	Denmark, Spain, Sweden	Belgium, Denmark, Finland, France, Poland, Sweden
Mental health and new psychosocial risks	Sweden		Belgium, Finland, France, Germany, Netherlands, Poland, Spain, Sweden
Threats to collective bargaining			France, Spain

Source: *Authors' elaboration, based on interviews with policy stakeholders*

National policy initiatives

In the last five years, new national policy initiatives have emerged bringing into focus ethical issues arising from the use of new digital technologies, particularly AI. This greater focus on ethics in national policy initiatives mirrors the increasing 'ethification' of EU policies and discourse on AI.

A comparatively higher number of policy initiatives has been launched in countries where debates on ethics are an extension of earlier and longer standing discussions on the implications of Industry 4.0 technologies for work and employment. This is, for example, the case in Denmark, France, Germany and Spain. The workplace dimension remains, however, largely overlooked in policies, with national initiatives drawing attention to the broader ethical implications of AI for society.

Government-led initiatives

National strategies on digitalisation are typically broad in scope, aiming to promote the uptake of new technologies in businesses and society at large, to increase competitiveness and enhance public service delivery. The launch of the Recovery and Resilience Facility (RRF) and the creation of national recovery and resilience plans have accelerated the adoption of AI plans and a range of policy initiatives on the digital transformation. Beyond a general requirement to contribute to digital transformation, the RRF requires EU Member States to dedicate at least 20% of their total allocation to measures contributing to the digital transition or to addressing the challenges resulting from it. The RRF supports reforms and investments to promote the roll-out of high-capacity networks, the digitalisation of public services and government processes, the digitalisation of businesses and the development of digital skills, as well as measures supporting research and development related to digitalisation and the deployment of advanced technologies.

With regard to the implications of new technologies for working conditions, a recurrent theme in the recovery and resilience plans and national digitalisation or AI strategies is the importance of upgrading skills in an increasingly digitally driven society. The commitment is invariably not limited to strengthening formal education in AI and digital skills, but includes fostering a lifelong learning culture and creating continuous upskilling and reskilling opportunities for the workforce, with a focus on digital skills. In spite of the emphasis on education and lifelong learning in national AI strategies, only a few of these strategies report quantified investments in this area (with the most notable examples being Denmark, Finland, France and the Netherlands) (Foffano et al, 2023).

The emphasis on ethics is more apparent in national AI strategies than in broad national digitalisation strategies. The latter generally focus on the benefits of digitalisation, in terms of economic growth and innovation. However, in some countries, national digital strategies and plans also refer to ethical considerations. For example, in the Netherlands, the National Digitalisation Strategy seeks to facilitate dialogue, support research and raise awareness on ethical issues through various initiatives, such as the development of a toolbox for ethically responsible innovation and an ethical code for good digital governance in public institutions. More specifically with regard to AI, the Dutch Ministry of the Interior and Kingdom Relations also commissioned several research teams (Vrije Universiteit Brussel, Tilburg University, Eindhoven University of Technology and the Netherlands Institute for Human Rights) to develop guidelines aiming to correct discriminatory algorithms and setting out the

technical, legal and organisational requirements for the development and use of AI applications in both public and private organisations. As part of the Dutch national AI strategy, resources have also been allocated to studying the impact of AI on work and employment.

Ethical considerations are also at the heart of the Danish national strategy on digitalisation, which is underpinned by nine 'visions' or ambitions, one of which refers to 'a strong, ethical, and responsible digital foundation' (Danish Government, 2022). The strategy draws attention to the ongoing need 'to address and discuss ethical issues and dilemmas' (p. 68). To this end, as part of the implementation of the strategy, the Danish Data Ethics Council, established in 2019, is set to continue its work by providing a forum for discussing ethical issues arising from the use of new technologies and designing tools for companies, such as a data ethics toolbox, to help them use data ethically and responsibly.

A similar consultative body was established in 2019 by the French government under the aegis of the National Consultative Ethics Committee and known as the National Pilot Committee for Digital Ethics (Comité National Pilote d'Ethique du Numérique). The committee is tasked with identifying ethical issues arising from the use of new technologies and issuing recommendations to both government and industry for the ethical use of such technologies. For example, a 2021 opinion released by the committee concerned the ethical implications of human–machine interactions through chatbots (for instance, deception occurring when humans are led to believe they are interacting with another human being and not a machine) (CNPEN, 2021).

LaborIA is another national initiative, launched in 2021 by the French Ministry of Labour, Employment and Economic Inclusion (in collaboration with the National Institute for Research in Digital Science and Technology). The initiative, part of the National AI Research Programme (part of the national AI strategy), involves setting up a research laboratory tasked with investigating the effects of AI technologies on work, employment, skills and social dialogue (Inria, 2022). Set to operate for five years, the laboratory will conduct in-company experiments as well as a barometer based on a survey of 250 companies to understand the impact of AI technologies on companies and workers (Actu IA, 2022).

A range of similar initiatives were launched in Germany by the Policy Lab Digital, Work & Society (Denkfabrik Digitale Arbeitswelt), established in 2018 by the Federal Ministry of Labour and Social Affairs and serving as an interdisciplinary think tank, to support the implementation of the national AI strategy. Examples of projects implemented by the Policy Lab Digital, Work & Society are the Observatory for Artificial Intelligence in Work and Society and the Employee Data Protection Committee. While the former has a broad focus, dealing with both the ethical implications of AI use and its implications for working conditions, the latter is specifically dedicated to data protection and privacy issues arising from the use of digital technologies in the workplace.

Also established in 2018, the Enquete Commission on AI is responsible for providing insight on the development of AI and its broad implications, including for employment and work, and ultimately informing the implementation of the national AI strategy.

Particularly proactive is the policy approach taken by the Spanish government, as exemplified by a series of interlinked initiatives aimed at supporting the future implementation of the AI Act, with a strong emphasis on establishing guarantees for the protection of fundamental human rights. Most of these initiatives are part of the implementation of the national AI strategy and include, for example:

- the launch of a regulatory testing environment ('sandbox') to study the feasibility of the requirements set out in the AI Act and develop guidelines and tools for the ethical and lawful development of high-risk AI systems[11]
- the development of a trustworthy AI seal (as a quality label)[12] for non-high-risk AI products and compatible with the European regulatory framework
- the creation of a national agency for the supervision of AI (through Law 22/2021 of 28 December on the general state budget for 2022), responsible for developing, supervising and monitoring AI-related projects
- the establishment of an AI advisory council providing recommendations to the government on actions to ensure the safe and ethical use of AI in society

11 While the initiative was launched in 2022, the Spanish government published a draft royal decree in June 2023, providing the legal basis for the development of the AI sandbox. The royal decree will be valid for 36 months after its entry into force and will expire as soon as the AI Act becomes applicable (Diario La Ley, 2023).

12 Some experimentation is ongoing in EU Member States on the development of ethical labels for AI products and services to ensure that ethical criteria are upheld and embedded in their design and development. These efforts are often led by research institutes and non-profit organisations active on AI issues and bringing together the industry and the research community. See, for example, information on the quality seal developed by the German AI association at https://www.hannovermesse.de/en/news/news-articles/quality-seal-for-ai-to-reinforce-ethical-values

The urgency of developing an ethical and legal framework for AI technologies is also emphasised in the national digital strategy (measure 45 of the Digital Spain 2025 agenda) (Ministerio de Asuntos Económicos y Transformación Digital, 2019).

More specifically with regard to algorithm auditing, the Spanish government also established in 2021 the Observatory on the Social and Ethical Impact of Algorithms (Observatorio del impacto social y ético de los algoritmos) to develop a methodology for auditing algorithms and reference indicators for the audits. Not specifically focused on AI, the Charter of Digital Rights (*Carta de derechos digitales*), with a section on digital labour rights adopted by the government in 2021 (Government of Spain, 2021), is another milestone initiative serving as a reference framework and a guide for future legislative projects and policies.

Although the national AI strategies in Finland and Sweden recognise the importance of the human-centric and trustworthy development and ethical and safe use of AI, fewer government-led initiatives in these countries explicitly focus on ethics, and even fewer are related to ethics in the context of work and employment. In Finland, a major initiative related to ethics was the establishment of an ethics board in 2020 as part of the national AI programme AuroraAI. The board is tasked with identifying ethical pitfalls in AI development and proposing ethically sustainable and human-centric public service solutions (Leikas et al, 2022). Broader in scope is the Committee for Technological Innovation and Ethics, supporting the Swedish government in identifying policy and regulatory challenges and accelerating policy developments linked to Industry 4.0 technologies.

In Poland, there are no noticeable initiatives referring to or addressing ethical challenges posed by the digitalisation of work. The discussions within working groups set up by the government in the context of the development of the national AI strategy (adopted in 2021) are very broad in scope and only marginally focus on the ethical side of technology development and use, particularly in work settings.

Multistakeholder initiatives

Examples of broad multistakeholder initiatives with a focus on ethics – either initiated by governments or where the government plays an important role – can be found in a handful of countries, most notably in Belgium, Germany, the Netherlands and Sweden.

An example is the AI4Belgium coalition, initiated in 2019 by the Belgian government (the two Ministers for the Digital Agenda), seeking to create a bridge between different stakeholders from the public and private sectors and academia and civil society organisations. Setting up a new learning deal and the development of a responsible data strategy are among the key recommendations formulated by the coalition.

Markedly research-focused is the FARI institute, launched in 2021 by the Brussels Region as yet another partnership bringing together academic research, civil society and public administration with the aim of promoting research on human-centric, trustworthy, transparent and explainable AI across different areas.

Another national example of public–private collaboration is the Dutch AI Coalition, initiated by the Dutch government and bringing together several stakeholders from industry, educational and research institutions, and civil society organisations, with a view to supporting the development and use of AI across different areas. On the ethical front, the coalition has so far developed the ethical, legal and societal aspects concept to promote a collaborative, responsible and human-centric approach to AI development.

A similar initiative in Sweden is the AI Agenda for Sweden (AI-agendan för Sverige), launched by RISE, a state-owned research institute promoting innovation and research and performing testing and certification. The agenda was created in cooperation with social partners – including the Swedish Trade Union Confederation (Landsorganisationen) and the Swedish Municipal Workers' Union (Kommunal) – as well as other institutional actors, such as the Agency for Digital Government and some large companies. Non-discrimination, privacy, integrity and responsibility are important themes in the agenda, albeit not explicitly mentioned in relation to the workplace. The agenda nonetheless refers to the work environment as an area that ought to be further researched.

There are also examples of multistakeholder initiatives in Germany. These are mainly government led and are more explicitly focused on the work environment. The most notable examples are the taskforce on AI in the world of work (initiated by the Federal Ministry of Labour and Social Affairs) and the Economy 4.0 platform (initiated by the Federal Ministry of Education and Research and the Federal Ministry for Economic Affairs and Climate Action). These initiatives bring together employer organisations, trade unions, academics and research institutes, and government representatives to discuss the implications of technology use for work, develop common positions and provide guidelines on a wide array of topics of ethical relevance. Similar initiatives have been launched at regional level, such as the Economy & Work 4.0 platform in the North Rhine-Westphalia region.

Key legal frameworks and regulatory developments

As the AI Act is in the trilogue phase of the legislative process, most EU Member States have taken a 'wait and see' approach to regulating AI use in the workplace. New regulatory initiatives have nonetheless emerged over the past few years directly or indirectly addressing the ethical implications of new workplace technologies,

and building on the GDPR. Taken together, such initiatives signal a political willingness to deal with the disruptions brought by new digital technologies to work and employment.

Data protection and privacy law

New national regulatory initiatives – addressing ethical challenges arising from technology use – largely relate to data protection and privacy and build on core requirements set out in the GDPR.

The opening clause[13] under Article 88 of the GDPR provides that

> *Member States may, by law or by collective agreements, provide for more specific rules to ensure the protection of the rights and freedoms in respect of the processing of employees' personal data in the employment context.*
>
> (GDPR, Article 88(1))

Several EU Member States have, albeit to different extents, resorted to using the above opening clause to regulate employee data processing. In most of the countries investigated, national data protection legislation has provisions relevant to the use of digital technologies, including in the workplace. These are often complemented by provisions in labour law and other relevant legislation.

In Spain, Organic Law 3/2018 of 5 December on personal data protection and guarantee of digital rights (*Ley Orgánica de Protección de Datos Personales y garantía de los derechos digitales*) includes provisions regulating the right to disconnect in the workplace, the right to privacy in relation to the use of geolocation devices at work and the right to transparency with regard to the automated processing of personal data (Articles 88, 89 and 90). In recent years, the Spanish Data Protection Agency has also issued several guidelines on GDPR requirements in AI use, including a guide for adapting products and services using AI to ensure their compliance with the GDPR (AEPD, 2020) and a guide providing requirements for audits of data processing involving AI (AEPD, 2021).

In addition, in France, the national law on data protection (*La loi informatique et libertés*) – complemented by the decisions of the French Data Protection Authority (Commission Nationale Informatique et des Libertés) – provides a solid framework regulating the processing and use of personal data, explicitly prohibiting decision-making by machine alone. Since 2019, the data protection authority has issued several documents on high-risk applications such as biometric devices at work. These include, for example, the Model Regulation, which details the processing of employees' biometric data for controlling access to workplaces; new guidelines on the management of personal data related to HR activities; and a deliberation on the use of biometric devices for access to premises, which lists high-risk data processing operations involving biometric devices (including for work monitoring and human resource management) and hence requiring a data protection impact assessment.

Similarly, in Finland, the Act on the Protection of Privacy in Working Life (*Työelämän tietosuojalaki*) provides the main legal framework for safeguarding the protection of privacy in working life and other fundamental rights, including in the context of technological workplace surveillance. In Sweden, data protection legislation is complemented by specific provisions on workplace surveillance in the Work Environment Act (*Arbetsmiljölagen*) and the regulation on work in front of a screen *(Arbete vid bildskärm*). The latter bans screen-based work when 'closely controlled or restricted in a physical or mental respect or [when it] is monotonously repetitive'. Recommendations and information on computer work issued by the Swedish Work Environment Authority (undated) draw attention to the negative effects of computer-based surveillance on both physical and mental health. Due to the pervasive nature of digital technologies, the Swedish Work Environment Authority has already for some time been calling for the modernisation of the legislative framework governing the digital work environment (2015).

In Germany, data protection authorities have called for a specific data protection law applicable to the workplace context. There are 17 state data protection authorities: one for each of the 16 federal states, with Bavaria having two (one for private organisations and the other for public organisations). According to a 2022 report prepared by an interdisciplinary advisory board appointed by the Federal Ministry of Labour and Social Affairs, an employee-specific data protection legislation is much needed. The creation of such a piece of legislation now features high on the policy agenda of the German governing coalition (for more information, see SPD et al, 2021). In addition, the German Trade Union Confederation (Deutscher Gewerksschaftbund, DGB) published in February 2022 the draft Employee Data Protection Act, addressing several ethical issues arising from the use of AI systems and intrusive surveillance technologies in the workplace (DGB, 2022).

13 Opening clauses are commonly used in the harmonisation of EU law to facilitate political agreement among EU Member States over contentious issues in the legislative process.

Another national initiative relevant to data protection is the establishment in the Netherlands of a new algorithmic supervision body reporting to the Dutch Data Protection Authority (IAPP, 2023). Core tasks of this supervisory body are the identification and analysis of the cross-sectoral risks and effects of algorithms (including discrimination, exclusion and a lack of transparency) and sharing knowledge with other relevant authorities.

In addition, a host of decisions and fines are issued by national data protection authorities over the infringement of GDPR rules in the context of employment. One of the most exemplary is the €35.3 million fine imposed by the Hamburg Commissioner for Data Protection and Freedom of Information on the H&M Service Center in Nuremberg for the illegal surveillance of hundreds of employees (EDPB, 2020). A lower, yet significant, fine (€20 million) was imposed by the French Data Protection Authority on software company Clearview AI over its facial recognition software. The decision revealed that the software presented significant risks to fundamental rights resulting from the unlawful collection and processing of personal data (EDPB, 2022). Smaller fines have been issued in recent years by national data protection authorities to companies for the use of intrusive processes for electronically monitoring workers, in violation of GDPR rules (GPDP, 2021; TechGDPR, 2022).

Labour law and other relevant legislation

New provisions have been introduced in recent times in national labour legislation to deal with new regulatory challenges posed by digital technologies in the context of work. One notable example is the newly introduced Article 64.4 in the Spanish Workers' Statute (*Estatuto de los Trabajadores*), providing that worker representatives must be 'informed by the company of the parameters, rules and instructions on which the algorithms or AI systems are based that affect decision-making and that may affect working conditions, access and maintenance of employment, including profiling of workers and applicants'. This provision applies to all companies using algorithmic management and not only platform companies operating in the food delivery sector.

In Sweden, paragraphs related to digitalisation have been added in recent years to the Work Environment Act, which regulates work environment activities and provides the basis for all occupational health laws. For example, the act requires that the work environment takes into account technological developments in society and establishes that employers are required to provide employees with the new skills needed as a result, on a non-discriminatory basis – that is, without leaving any employee behind in the digital transition. In addition, the 2021–2025 Work Environment Strategy notes that skills development is a shared interest of both employers and employees and that in situations when jobs are assigned using novel digital tools, it should be clear who is responsible for health and safety at work (Government of Sweden, 2021). The strategy also states that it is essential that 'working conditions continue to be predictable and safe, even in a world that is changing' (Government of Sweden, 2021, p. 13).

Another recent regulatory development concerns Denmark, where in May 2020 the parliament adopted an amendment to the Danish Financial Statements Act (*årsregnskabsloven*). This requires large companies that have a policy for data ethics to supplement their management's reports with an account of the company's policy on data ethics. The aim of the legislation is to effectively compel companies to create and define data ethics statements that incorporate transparency, explainability and bias mitigation into their brand.

Building on Directive 2002/14/EC of 11 March 2002 on informing and consulting employees, national legislation in countries with a strong tradition of social dialogue grants worker representatives the right to be informed and consulted about significant changes to work organisation and working conditions, making explicit reference to technological innovation. An example is the German Works Constitution Act, conferring on works councils co-determination rights in the introduction and use of technical devices to monitor the behaviour or performance of employees. The act provides that the use of technical devices must be necessary and proportionate. This assessment is to be made on a case-by-case basis. Monitoring is not always deemed unlawful. For example, in February 2023, the Hanover Administrative Court ruled in favour of the Amazon logistics centre in Winsen monitoring the working speed of its employees with the help of hand-held scanners (Case No. 10 A 6199/20); in the court's opinion, the continuous collection and processing of data on employees' performance were deemed necessary for the implementation of the employment relationship and were in compliance with data protection regulations.

A complementary piece of legislation is the Works Council Modernisation Act, which came into force on 18 June 2021 and extended the co-determination rights of works councils with regard to the introduction and operation of AI systems in the workplace. Although the DGB welcomed this new legislation, the confederation also expressed concerns that it is not far reaching enough; for example, it does not establish a binding process for introducing AI systems (DGB, 2021).

In addition, in other countries, the employer is obliged by law to consult the works council and/or other employee representatives over the introduction of new workplace technologies if these are deemed to affect working and employment conditions (see Table 9).

Table 9: Examples of national legislation on employee representatives' rights to information and consultation in relation to the impact of digitalisation

Country	Key legal frameworks
Denmark	The Information and Consultation of Employees Act provides that the employer must consult with employee representatives on changes affecting, for example, employees' salaries, work descriptions, working hours and location of work. This includes changes resulting from the introduction of new technologies in the workplace. Statutory health and safety regulations also require employers to inform and consult employees about the introduction of new technologies and any implications they may have for working conditions and work organisation. Employers must also involve employees in deciding on ways to deal with these changes, including through updated training requirements.
Finland	The Act on Co-operation within Undertakings (334/2007), the Act on Cooperation within Government Agencies and Institutions (1233/2013), and the Act on Cooperation within Municipalities (449/2007) provide the legal framework for negotiations between employers and employees in the workplace and set out a requirement for negotiations when introducing, for instance, new technologies. The legal framework includes provisions regarding the surveillance of employees through technical means. The Occupational Safety and Health Act (736/2002) does not mention technological change per se, but Section 17 (on cooperation between employers and employees) establishes that employers should in good time give the employees necessary information on any factors that affect health and safety in the workplace, and other circumstances that have an effect on their working conditions.
France	The Labour Code grants social and economic committees information and consultation rights over the use of the methods or techniques for recruitment purposes (and any subsequent modifications), the introduction of automated processes for personnel management (and of any subsequent modifications) and the decision to introduce means or techniques to control employees' activities (Article L.2312-38).
Germany	The Works Constitution Act establishes co-determination rights in the introduction and use of technical devices designed to monitor the behaviour or performance of employees and requires the employer to inform the works council in a timely fashion of any plans concerning working procedures or operations. The Works Council Modernisation Act extends the co-determination rights of works councils to decisions regarding the introduction and use of AI in the operations of companies, including for HR purposes. It also provides that works councils can call on external ICT experts if needed.
Netherlands	The Dutch Works Council Act of 2018 provides works councils with a right to be consulted about, among other things, the introduction of new technologies or the alteration of existing technologies used in the workplace. The works councils should also be consulted about any changes that have an impact on the work environment. The employer must delay taking action for at least a month if the works council disagrees with its proposal. During this time, the works council has the right to appeal to the Companies Chamber of the Court of Appeal in Amsterdam.
Spain	Royal Decree-Law 9/2021 of 11 May, amending the revised text of the Workers' Statute, provides that worker representatives must be informed by the employer of the parameters, rules and instructions on which the algorithms or AI systems are based, that affect decision-making and that may affect working conditions and access to and the maintenance of employment, including profiling of workers and applicants.

Source: *Desk research and interviews with policy stakeholders in the selected countries*

As part of a policy study on the use of AI and digital tools in the workplace, the European Parliamentary Research Service pointed to the uneven and 'somewhat' unsatisfactory implementation of Directive 2002/14/EC and called for a review of EU Member States' national laws to best protect worker representatives' information and consultation rights in the digital age (EPRS, 2022a).

Another issue is that technologies are very versatile and can be easily redeployed for other purposes beyond those initially specified and communicated to employees; it is therefore important to ensure that employers re-engage in negotiations with employee representatives in situations where technologies are used for a different purpose or scaled up to ensure that quality of work is safeguarded and that workers' rights are not compromised.

Self-regulation through collective bargaining

Collective bargaining is a complex practice that requires time to adapt to changing realities, and this may be the reason for the still limited coverage of digitalisation-related issues in collective bargaining. Based on a mapping and evaluation of collective bargaining provisions on workers' rights in the context of digitalisation, Voss and Bertossa (2022) concluded that 'digitalisation is still not automatically or comprehensively included in collective bargaining agendas' and 'most provisions cover areas of traditional union work or areas where existing models can be adapted' (p. 29).

There are, however, instances where national social partners have negotiated new provisions in collective agreements in an effort to regulate the use of digital technologies in the workplace. An example of a forward-looking cross-industry agreement is the employment and collective bargaining agreement (*V Acuerdo Colectivo para el empleo y la negociación colectiva*), reached in 2023 by the peak-level Spanish social partners. This agreement formally implements the EU social partners' autonomous framework agreement on digitalisation (ETUC et al, 2020) and includes one chapter that addresses issues arising from the use of algorithms and AI systems in companies,

particularly in HR procedures. Other contemporary collective agreements include provisions that address issues surrounding new skills needs, privacy and data protection, particularly in the context of monitoring employees through digital means, as well as procedures for connecting and disconnecting from work, which affect employees' work–life balance and well-being.

Belgium is one of the countries regulating issues around privacy, monitoring and data protection through national collective agreements. These include the long standing collective agreement No. 39 of 13 December 1983 on information and consultation with regard to the social consequences of the introduction of new technologies, collective agreement No. 81 of 26 April 2002 on the protection of the private life of workers with regard to the monitoring of networked electronic communication data and collective agreement No. 68 of 16 June 1998 on the use of camera surveillance in the workplace.

In other countries, some sectors and companies have negotiated collective agreements that set frameworks for responsible workplace digitalisation. Innovative in this respect is the agreement reached in March 2021 by the Spanish Banking Employers' Association (Asociación Española de Banca) and its associated trade unions – the Trade Union Confederation of Workers' Commissions (Comisiones Obreras, CC.OO), the General Union of Workers (Unión General de Trabajadores, UGT) and the Independent Trade Union Federation (Federación Fuerza, Independencia y Empleo). The collective agreement recognises a set of digital rights, including the right to disconnect, the right to digital privacy, the right to digital education and the right to not be subject to fully automated algorithmic decisions, and not be discriminated against based on such decisions.

In other countries, such as Belgium, Germany and Italy, collective agreements in the area of banking cover a narrower set of digitalisation-related issues often linked to new skills requirements and the need to prepare the workforce for the digital transition (Eurofound, 2021b). For example, in Italy, the collective agreement in the banking sector – signed in December 2019 and renewed in 2022 – establishes a bilateral and joint national committee tasked with monitoring the digital transformation in the sector and identifying new skills needs. Similarly, in France, the 2020 national collective bargaining agreement for the banking sector established an observatory of professions, qualifications and professional equality. The observatory is tasked with monitoring the quantitative and qualitative evolution of employment and professions, identifying changes due to technological developments, and providing information to social partners and companies to facilitate social dialogue on digital issues (AFB et al, 2020).

However, when it comes to digital technologies, sectoral collective agreements tend to lack binding provisions and instead foster opportunities to further discussions and social dialogue on the digitalisation of work. For example, the Spanish collective agreement covering large retail companies provides for the establishment of a sectoral observatory that functions as a forum for social dialogue to analyse developments in the sector, with a particular focus on changes triggered by digitalisation (Eurofound, 2022b). In Finland, the 2022–2024 collective agreement for the finance sector envisages the continuation of discussions on the future of work as part of the FinanssiTYÖ 2030 project. Similarly, the 2022–2024 collective agreement for the ICT sector establishes a specific working group on the impact of digitalisation, robotics and AI on work.

The collective agreement for the metal industry concluded in Germany in 2021 by the metalworkers' union IG Metall provides the possibility of negotiating company-specific 'future-oriented collective agreements' (*Zukunftstarifverträge*). These may introduce new rules not included or diverging from the industry collective agreement, depending on the particular characteristics of companies. While they are generally intended to be used in restructuring situations, they may become more important in the context of technology-induced redundancies.

In Germany, where collective bargaining also often takes place at company level, some company-level collective agreements deal with the impact of digitalisation on work. One is the collective agreement concluded in 2019 between the German service sector trade union Vereinte Dienstleistungsgewerkschaft (ver.di) and the port operator Eurogate. The agreement runs for 10 years and prevents dismissals for operational reasons until 2025. In addition to the employment guarantee, the operator and the trade union also agreed on new working time models and means of expanding employees' qualifications. With regard to the latter, employees at Eurogate have a right to have their career prospects assessed within the company and train to obtain further qualifications according to their needs. Employees who are no longer needed in their current occupation due to automation are entitled to be offered another job within the company.

Another example of a company-level agreement is that signed by ver.di and the German subsidiary of the Swedish clothing retailer H&M, which introduces a range of measures mitigating the impact of digitalisation on the workforce (Mind, 2022). These include employment protection guarantees, requalification and redeployment measures, and the creation of a digitalisation committee. The company also commits to involving employees in the company's digital strategy and the development of 'a concept for the future'. Other company-level agreements – for

example, those of aircraft manufacturer Airbus and pharmaceutical company Merck – concluded with works councils seek to establish new forms of strategic cooperation and social dialogue to deal with the challenges of digitalisation (Krzywdzinski et al, 2023).

With a more specific focus on AI systems, IBM Germany and the works council developed a framework agreement that defines key standards and principles in relation to AI use; these cover, for example, human in control, transparency, explainability, non-discrimination and quality assurance. Under this agreement, the use of AI systems that automate decisions about humans are explicitly prohibited, and an AI ethics council was established to monitor developments (Krzywdzinski et al, 2023).

Other company-level agreements involve the creation of instruments for monitoring change. One example is the collective agreement signed in 2020 by textile group Inditex (with 48,000 staff in Spain alone) and the trade unions CC.OO and UGT. The agreement established a joint observatory for monitoring the process of digital transformation in Inditex stores and more generally managing the transition to an e-commerce business model.

At international level, the global framework agreement on digital transformation concluded by Belgian multinational chemical company Solvay and its European works council, supported by IndustriAll Global Union, is one example of an agreement that is part of the new wave of collective agreements on managing the changes driven by the increasing digitalisation of work. The agreement establishes 'a framework and a set of principles' for local management and employee representatives to address the issues that arise from the digital transformation in detailed joint action plans. There are no references to specific technologies in the agreement. Digital technologies are rather referred to in general terms, as 'electronic tools, systems, devices and resources that generate, store, process, exchange or use digital data'. Issues that feature more prominently in the agreement relate to privacy, data protection, and implications of working in a digitally connected environment for employees' work–life balance and mental health.

Although issues related to digitalisation – particularly of an ethical nature (data ethics, algorithmic management, worker surveillance and AI governance more generally) – are not systematically covered in collective bargaining, the abovementioned initiatives demonstrate the added value of collective bargaining in dealing with emerging challenges in the interests of both workers and employers. This is done by safeguarding existing workers' rights and/or establishing new digital rights and laying the foundation for socially acceptable and new ethical working practices that are fit for the digital age.

Social partners' role

Social partners' involvement in government initiatives

The involvement of social partners in government-led initiatives varies across the nine EU Member States investigated. In countries with a strong culture of tripartite social dialogue, social partners are represented in various ways. For instance, in Belgium, the National Labour Council (Nationale Arbeidsraad/Conseil National du Travail), which represents major employer associations and trade unions, is consulted or provides opinions unprompted on general issues of a social nature, including the social consequences of the introduction of new technologies. The council's opinion No. 125 of 6 April 2022 aligned the planned and implemented actions supporting the digital transition with the broad objectives set out in the European social partners' framework agreement on digitalisation.

Although France, Germany and the Netherlands are among the countries where social partners are involved in various government-led initiatives, the social partner representatives consulted as part of this study – particularly trade unions – indicated the need to further strengthen tripartite social dialogue on digitalisation-related issues. In this regard, French trade union representatives argued that the role of social partners is mostly consultative and called for their greater involvement upstream in the policymaking process.

Concerns were also expressed by the Spanish social partner representatives, who denounced the lack of consistent tripartite social dialogue on issues related to the digitalisation of work. In Poland, the role of social partners has so far been limited to participating in discussions within the Social Dialogue Council, where issues surrounding digitalisation of work are also being debated (mainly pertaining to data protection and privacy, the right to disconnect and access to training to improve digital skills).

In Finland, social partners are traditionally involved in designing labour market policies. For example, the 2019 amendment of the Act on the Protection of Privacy in Working Life was prepared by a tripartite working group. With regard to the AuroraAI programme, the role of the Finnish social partners was instead rather consultative. According to the Finnish trade union representatives interviewed, the social partners should have had a more significant role in the design of this programme.

Social partners' initiatives

European social partners

Since the adoption in 2020 of the European social partners' framework agreement on digitalisation, national social partners have been negotiating and implementing measures and initiatives mainly in two of the four areas of focus in the agreement – that is, 'digital skills and securing employment' and 'modalities of connecting and disconnecting' (ETUC et al, 2021, 2022). These topics were regarded by the social partners as urgent priorities in the face of the persistent skills and labour shortages experienced in many countries and the massive surge in remote working in the aftermath of the pandemic. Initiatives addressing the ethical implications of digital workplace technologies (in the areas entitled 'AI and guaranteeing the human in control principle' and 'respect of human dignity and surveillance') remained somewhat overlooked in social partners' actions. In addition, issues around skills and working time have been dealt with systematically in law and collective agreements already for some time, while social partners are still getting to grips with the more complex and technical issues related to AI and algorithms and their implications for work.

In some instances, sectoral European social partners have issued joint declarations emphasising the key role of social dialogue in dealing with the challenges posed by the digitalisation of work. An example is the joint position paper of the European Tech and Industry Employers (Ceemet) and the European trade union federation IndustriAll Europe on the impact of digitalisation on the world of work in the metal, engineering and technology-based industries, signed in 2020 (Ceemet and IndustriAll Europe, 2020). The document identifies four topics for social dialogue to ensure 'the best possible outcomes for both employers and workers' in the ongoing digital transformation of the industry; these are work organisation, skills, health and safety at work, and data protection. The importance of social dialogue for dealing with the impacts of new technologies on employment is reiterated in the joint conclusions of Ceemet and IndustriAll Europe on AI in the metal industries, adopted in February 2023 (Ceemet and IndustriAll Europe, 2023).

A joint declaration on AI was also signed in 2020 by telecom employer organisation the European Telecommunications Network Operators' Association and UNI Europa ICTS. The document emphasises the importance of transparency for building trust in technologies, setting up accountability mechanisms, and ensuring a human-centric approach to technology design and implementation. In the joint declaration, the signatory parties commit to developing an ethical framework laying the foundation for 'the use of AI to the benefit of both societal and economic objectives' (ETNO and UNI Europa ICTS, 2020). Similarly, in 2021, the European social partners in the insurance sector – UNI Europa Finance, Insurance Europe, BIPAR and AMICE – signed a joint declaration on AI whereby they commit to promoting the sustainable, responsible and ethical use of AI. According to the signatory parties, this means designing and using AI 'to enhance rather than replace human abilities' (UNI Europa, 2021).

In 2020, UNI Global Union also published a guide providing recommendations for trade union members on how to negotiate algorithmic management in the workplace (UNI Global Union, 2020). The focus is on ethical principles including transparency, accountability, proportionality, equity, access to data and human in control. In a 2022 survey conducted by UNI Europa and the Friedrich-Ebert-Stiftung's Competence Centre on the Future of Work, 34% of respondents (1,400 workers across the EU), predominantly workers in the ICT and telecommunications sector, reported that they did not know if algorithmic management tools were being implemented in their workplace (Friedrich-Ebert-Stiftung, 2022).

National social partners

In recent years, national and sectoral social partners have actively contributed to discussions and debates around the ethical implications of new digital technologies and their impact on working conditions by issuing position papers, recommendations and guidelines, as well as conducting ad hoc studies. A shared understanding of some of the challenges arising from the digitalisation of work is apparent from a set of recommendations signed by a group of Finnish social partners. These recommendations revolve around the need to invest in the continuous upskilling, reskilling and training of workers; establish coherent and fair rules; and adapt working practices to a digital work environment (Akava et al, 2019). The signatory parties also argue that cooperation between social partners will facilitate the adoption of AI, improving productivity as well as having broader benefits for workplaces in terms of workers' well-being and improving working conditions.

In most countries, social partners generally agree that there is a significant need for competency development and skills upgrading at work. The Spanish trade unions CC.OO and UGT and the employer association representing the digital technology industry, Ametic, published in 2020 a joint report highlighting the need to develop digital skills and ensure the adequate provision of training for the acquisition of new digital skills (CC.OO et al, 2020). A similar concern was expressed in a joint analysis conducted by the Confederation of Danish Industry and the transport and logistics department of the United Federation of Workers in Denmark in relation to the effects of rapid technological change on the transport sector (DI and 3F, 2019). On the matter of skills, the General Confederation of Liberal Trade Unions of Belgium has put forward some proposals for dealing with emerging skills gaps, calling for a

guaranteed right to vocational training, particularly for people at risk of technological unemployment, and the creation of sectoral digitalisation funds for training and reskilling (CGSLB, 2018).

In Sweden, job security councils – jointly owned by employer organisations and trade unions – are particularly active in supporting displaced workers and play an important role in the anticipation of skills needs (OECD, 2019). As evidenced by policy documents published by Swedish social partners, while there is general agreement on the need for skills and competency development for all workers (LO, 2016; SACO, 2017; TCO, 2018; Confederation of Swedish Enterprise, 2020a; Unionen, 2020), views diverge as to the funding of training and requalification efforts. For example, the Confederation of Swedish Enterprise (2020a) argues that funding must be provided by the state and the social partners, while the Swedish Trade Union Confederation believes that employers must bear part of the cost of competency development (LO, 2016). Social partners – namely the Confederation of Swedish Enterprise, the Negotiation and Cooperation Council, the union IF Metall and the Swedish Municipal Workers' Union – have come to an agreement in the context of updating the Swedish Employment Protection Act, which contains provisions relevant to lifelong learning. The new agreement allows staff to take time off work for competency development and could therefore assist in developing relevant skills for new tasks or new jobs. The agreement also increases employers' responsibility for funding investments in training and provides that employees who choose to develop their skills should receive funding through job security councils (Confederation of Swedish Enterprise, 2020b).[14] The amended legislation came into force on 30 June 2022 and promotes a structural shift in the Swedish model for skills development through the establishment of 'fundamental transition and skills support'. Workers who are not covered by a collective agreement can get support in the form of advice and guidance when they need to develop their skills to enhance their employability.

Trade unions' position papers and studies also call for a participatory approach to workplace digitalisation. In Belgium, the General Labour Federation of Belgium published a position paper in 2019 stressing the importance of employee participation in digitalisation processes, through the involvement of works councils or union representatives at an early stage of the process (ABVV-FGTB, 2019). At federal level, the report *Social partners on the digital fast track*, issued by the Social and Economic Council of Flanders (SERV, 2018), acknowledged the need for social partners to develop new approaches to and ideas for dealing with the emerging challenges posed by digitalisation and successfully manage the effects of technological developments.

Employees' participation in AI design and adoption is also central to the idea of 'good work by design' proposed by the DGB in a recent concept paper (DGB, 2020). According to the DGB, 'a prerequisite for good design is a broad participation process, which should begin with the definition of the objectives for the AI and its application and should include an impact assessment' (DGB, 2020). The DGB concept paper also identifies challenges arising from the increasing digitalisation of work. These include job losses due to increased automation; new occupational health risks, including those of a psychological nature; and unethical work surveillance and data handling (DGB, 2020). The idea of 'good work' is also central to the discussion paper and ethical guidelines issued by the German service sector trade union ver.di (2020) focusing on AI development and use in companies. The German metalworkers' union IG Metall has also launched several initiatives to help familiarise works councils with digitalisation-related topics (Transformation Atlas), assess the consequences of digitalisation for employees (Digitalisation Compass) and help employees develop the expertise required to deal with the issues arising (Work and Innovation 4.0) (Krzywdzinski et al, 2023). In addition, in other countries, such as Sweden, trade unions have been promoting initiatives to educate and train their members in AI and digital tools (TCO, 2022).

Several other trade unions have issued ethical guidelines and checklists for the responsible use of technologies in the workplace (Eurofound, 2022a). In most cases, guidelines published by social partners, albeit useful, do not provide practical guidance as to how to address ethical issues at each stage of the technology lifecycle. Furthermore, although the trade union representatives consulted regard ethical guidelines as being of added value, they consider that only national policies can truly operationalise them and achieve visible results in terms of the ethical, safe and responsible use of technologies in the workplace.

Calls for strengthening social dialogue in the context of the increasing digitalisation of work have come from both trade unions and employer organisations. In this regard, in Spain, the CEOE aims to broaden the debate on the future of work, engaging unions, the government, civil society organisations and professional associations (CEOE, 2020). In the context of the acceleration of the digital transformation prompted

14 Prior to this agreement, only employees who were about to be dismissed or were dismissed could access funding for training through the job security councils.

by the COVID-19 pandemic, the CEOE has also proposed some ways to address the risks of digital exclusion, for example by establishing a digital divide observatory and a catalogue of digital training plans for vulnerable groups, as well as promoting training for managers and workers on legal and security issues in the digital world of work (CEOE, 2020).

Social partners – especially trade unions – have been particularly active in conducting or commissioning studies on the challenges posed by digitalisation to the future of work, some of which are of an ethical nature. For example, a study commissioned by the Central Organisation of Finnish Trade Unions warns against some risks stemming from new-technology-driven trends such as the increase in employee monitoring and control, the greater datafication of work, the decrease in in-person social interactions and the erosion of solidarity within the workplace (Kesä, 2018). Employee surveillance is increasingly becoming a source of concern for some Swedish trade unions, particularly the white-collar trade union Unionen and the Confederation of Professional Employees (Futurion, 2020; Ingenjören, 2020; addAI, 2021).

From the employers' perspective, sectoral associations, most notably in Spain, have been channelling resources into monitoring AI use in companies and exchanging experiences among their members. Such efforts are exemplified by the Artificial Intelligence Observatory, launched by digital technology industry association Ametic to monitor the ethical and sustainable use of algorithms, particularly in the workplace; and the Observatory on Big Data, Artificial Intelligence and Data Analytics, created by the Spanish Accounting and Business Administration Association (Asociación Española de Contabilidad y Administración de Empresas) for the exchange of practices on the use of big data and AI and data analysis.

Key takeaways

- With the draft AI Act still under negotiation at EU level, most EU Member States have taken a 'wait and see' approach to implementing regulatory change. However, some countries – such as Spain and Germany – have implemented a range of policy initiatives aimed at addressing pressing ethical concerns arising from the digitalisation of work and, not least, preparing the ground for the effective national implementation of the upcoming regulation.
- Trust in technologies was a recurrent concern voiced by policy stakeholders. In the political discourse, the issue of trust is intertwined with considerations surrounding the accuracy, transparency and accountability of AI systems to avoid unfair or discriminatory outcomes. Public authorities have an important role in setting the ethical boundaries for the digital transformation of society, enacting policies that provide grounds for trust in technologies, engaging a broad range of stakeholders in the debate on ethics (including civil society) and, not least, leading by example in the ethical and trustworthy use of technologies for public service delivery.
- Although there is increasing awareness among stakeholders about the ethical implications of digital technologies for work and employment, most of the policy initiatives on ethics and digitalisation tend to be broad in focus, failing to sufficiently address the workplace dimension. The disruptive effects of digital technologies on work – with both ethical and legal ramifications – warrant more attention in policymaking and should be addressed in more workplace-focused policies. These could possibly be added to the catalogue of measures included in national digitalisation and AI strategies.
- Both stakeholder consultations and the review of policy documents suggest that digitalisation is not yet fully incorporated into collective bargaining. Contemporary collective agreements on managing the digital transformation typically address issues around working time and work–life balance (for example, in the context of remote working) as well as workers' reskilling and upskilling. These are traditional areas of concern (and action) for social partners. Sectoral and company-level collective agreements addressing ethical issues arising from technology (mis)use more comprehensively are the exception rather than the rule. This points to the need for social partners to exchange information, learn from each other and build on existing good practices.

3 Technology adoption: Survey and case study evidence

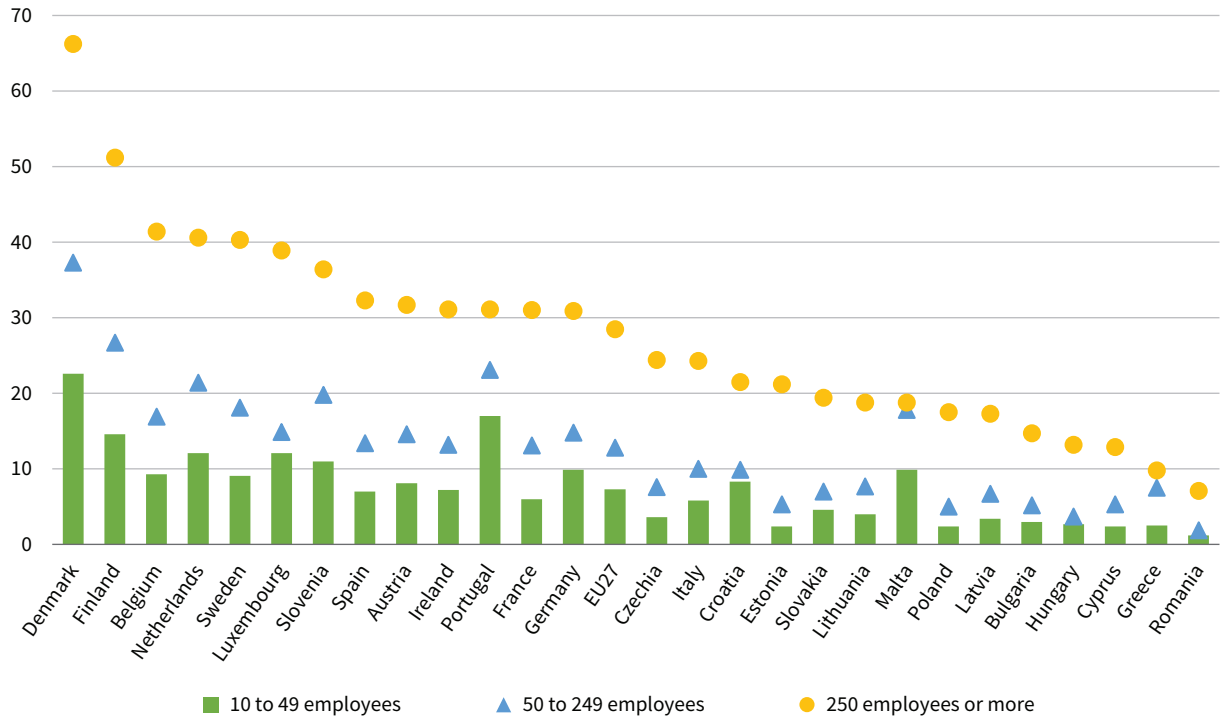
Adoption of artificial intelligence in enterprises

AI has progressively become one of the fastest growing technologies in the past decade in terms of both research and development and the range of applications that make use of its capabilities. Between 2013 and 2021, the volume of private investments in AI technologies increased 30-fold, with the US and China leading both in terms of total investments in AI and the number of newly funded companies (Zhang et al, 2022). Current forecasts predict that the impressive rate of increase in the use of AI witnessed since 2013 is going to continue over the next decade (PwC, 2020).

However, as of 2023, Europe remains a laggard in terms of the rates of AI investment and adoption. The most recent available data on the adoption of AI in the EU come from Eurostat's community survey on ICT usage and e-commerce in enterprises. The overall adoption of AI in the EU is low, with only 8% of enterprises with 10 employees or more having adopted at least one AI technology in 2021. However, the adoption rate varies significantly by enterprise size (Figure 2), with only 7.3% of small enterprises having adopted one type of AI system as of 2021. This share increases to 12.8% for enterprises employing between 50 and 249 employees and 28.5% for large enterprises, with 250 or more employees. The trends in adoption rates by enterprise size across countries largely replicate aggregate figures. AI adoption rates are higher in Denmark, Finland, Belgium, the Netherlands and Sweden, where more than 40% of large companies use AI systems. Conversely, the prevalence of AI use is lower in Bulgaria, Hungary, Cyprus, Greece and Romania, where fewer than 15% of large companies use at least one type of AI system.

The differences in adoption rates between large and smaller enterprises can be explained by the cost associated with implementing AI technologies in the workplace, the complexity of using AI systems and economies of scale, which allow larger companies to draw more benefits from investing and using AI technologies (Eurostat, 2023).

Figure 2: Share of enterprises using at least one AI technology in the EU, by establishment size (%)



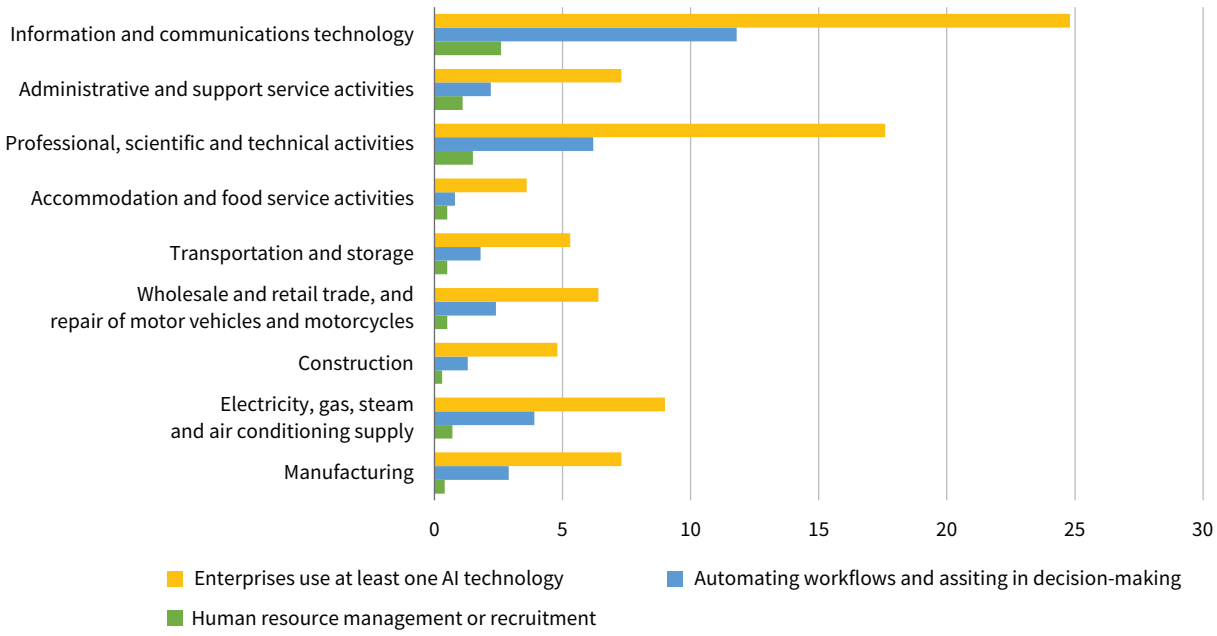
Source: *Authors' elaboration, based on Eurostat [isoc_eb_ai]*

In terms of sectors, AI is most common in the ICT sector, with nearly one-quarter of companies using at least one type of AI system (Figure 3). The use of AI is also relatively widespread in professional, scientific and technical activities (17.6% of companies), while it is less common in manufacturing (7.3% of companies), construction (4.8% of companies), and accommodation and food service activities (3.6% of companies).

The use of AI for automating workflows and assisting in decision-making is most common in the ICT sector, where it is present in 12% of companies.[15] This is most likely driven by the adoption of robotic process automation (RPA) systems, in which there has been a marked increase in recent years (McKinsey Global Institute, 2022). By comparison, the use of AI systems in human resource management or recruitment does not vary as much by sector or activity. Only 2.6% of companies in the ICT sector in the EU are using the technology for this purpose, compared with 1.1% of companies in the administrative and support service activities sector and fewer than 0.5% of companies in construction and manufacturing.

Based on Eurostat data, EU enterprises use AI technologies for a variety of purposes (Figure 4). AI systems are most commonly used by European enterprises for ICT security. This is the case for almost 40% of large enterprises and 20% of small enterprises that use at least one AI system. AI systems are also commonly used in production processes (for example, in predictive maintenance systems, cobot applications or inventory management systems) and in business administration processes (for example, in machine learning tools used for planning, smart assistants and automated responders providing online customer support). The least common uses of AI systems are in logistics and human resource management and recruitment. In the case of the latter, only 11% of large enterprises and about 8% of small enterprises employ an AI system for recruitment or human resource management purposes.

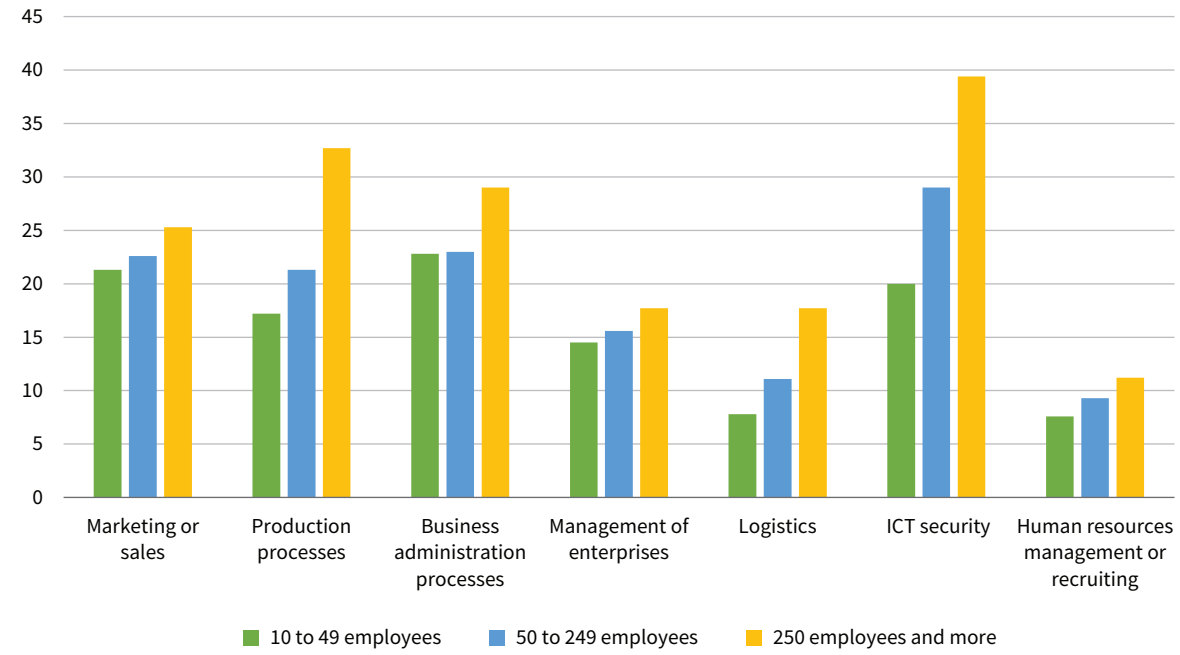
Figure 3: Share of enterprises using at least one AI technology in the EU, by sector of activity and purpose (%)



Source: *Authors' elaboration, based on Eurostat [isoc_eb_ain2]*

15 In this paragraph, quantities refer to the percentage of European enterprises using at least one AI technology.

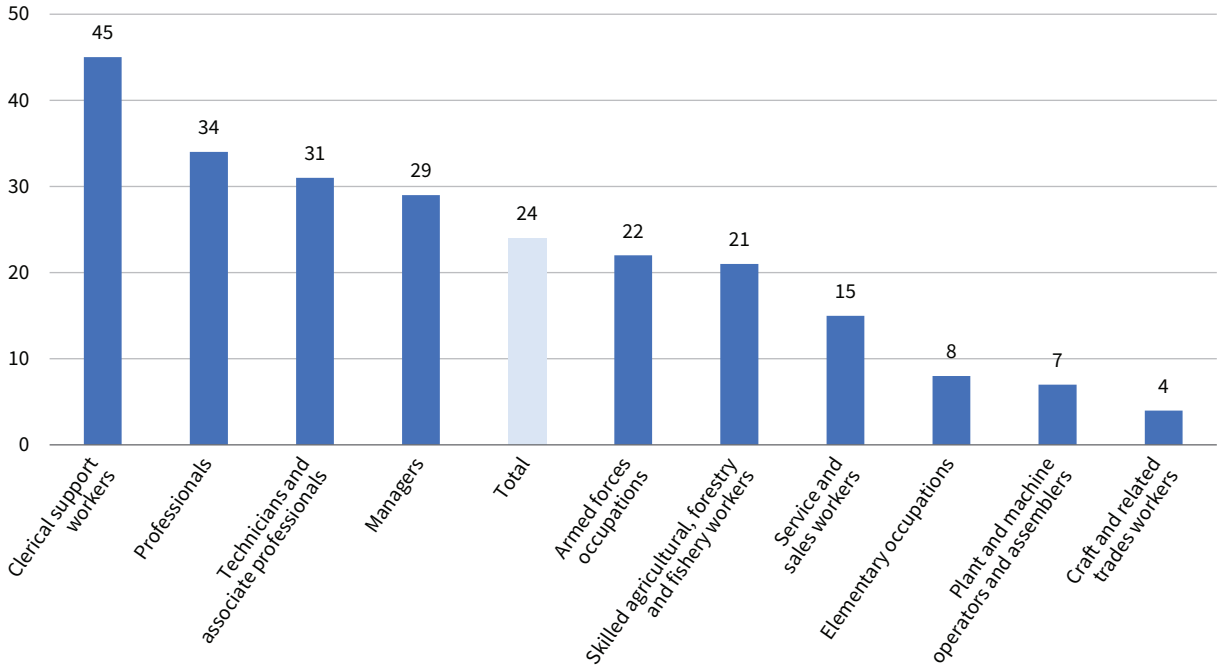
Figure 4: Share of enterprises using at least one AI technology in the EU, by size of enterprise and purpose (%)



Source: *Authors' elaboration, based on Eurostat [isoc_eb_ain2]*

Drawing on data on occupational tasks in the US and Europe, Goldman Sachs's research predicts that workers in about two-thirds of jobs are currently exposed to AI automation, with generative AI[16] substituting up to one-quarter of current work (Goldman Sachs, 2023; Figure 5). When extrapolating the data to the global level, the prediction is that about 300 million full-time jobs could be lost as a result of AI replacing workers. According to the investment bank research, white-collar occupations are more vulnerable

Figure 5: Share of industry employment exposed to automation by AI in the euro area (%)



Source: *Goldman Sachs, 2023*

16 The report defines generative AI technologies based on three main characteristics: (1) they have generalised rather than specialised uses; (2) they are able to generate novel, human-like outputs rather than merely describe or interpret existing information; and (3) they have approachable user interfaces that both understand questions and respond with natural language, images, audio and video.

to displacement by AI, whereas previous technological advances have been more likely to displace blue-collar occupations or more routine forms of work.

In an earlier study, the McKinsey Global Institute (2018) developed three different scenarios projecting the pace and scope of adoption of AI and automation technologies. In the midpoint scenario, about 15% of the global workforce (equal to 400 million workers) would be displaced due to automation by 2030. According to the same research, about 30% of activities in 60% of all occupations are technically automatable.

It is, however, not all doom and gloom, as automation is also expected to create innovation and ultimately lead to the creation of new jobs in the longer term (WEF, 2020; Goldman Sachs, 2023). Nevertheless, one important pre-condition for unlocking this positive trend is greater investment in workforce upskilling and reskilling (WEF, 2020, 2023).

Approaches to ethical digitalisation

Evidence from survey data

Available data sources suggest that ethical principles are often not embedded in technology development and use in the workplace. In a Deloitte global survey (2022) among 1,794 business and technical professionals (of whom 378 were in Europe), nearly 90% of respondents reported a lack of an ethical governance framework guiding technology development and use. An earlier Deloitte global survey (2020) identified four key drivers of the increasing importance of ethics in the future of work. These were legal and regulatory requirements, the rapid adoption of AI technologies in the workplace, changes in workforce composition and pressure from external stakeholders.

Another multi-country survey among 1,103 employers and 4,207 employees – in Australia, Germany, Japan, New Zealand, the United Kingdom and the US – on the effects of AI on their workplaces found that over half of surveyed employers indicated their company does not have a written policy on the ethical use of AI, and only 21% expressed concern that their companies could use AI unethically (Genesys, 2019). The survey also suggests that large companies are more likely to have a chief ethical officer (or a similar role), but they would generally deal with more traditional ethical issues in the workplace rather than those arising from the use of digital technologies. It is, however, worrying that, at the time of writing, large tech companies involved in the development of AI technologies – such as Microsoft, Meta, Google and Amazon – are downsizing their 'responsible AI teams' (The Guardian, 2023).

The McKinsey Global Survey on AI conducted online in 2021 among 1,843 executives suggests that awareness about ethical risks of AI (for example, explainability, privacy, equity and fairness) had increased in developed economies compared with the 2020 survey wave. However, there remains scope for improvement as to the management of such risks. In the survey, 30% of respondents in advanced economies reported that their organisation was working to mitigate risks related to explainability, followed by issues about privacy (29%) and equity and fairness (21%). Organisations classified as 'AI high performers'[17] are more likely than other organisations to engage in a range of risk mitigation practices, including model documentation, data validation and checks for biases.

The findings of national surveys also show that insufficient attention has been paid to ethical policies and guidelines in technology development and use. For example, in a survey conducted in Belgium among 620 business leaders (in companies of all sizes), about 70% of respondents reported that their company did not have any policy to ensure ethical responsibility in their use of AI tools and 33% conceded that it is likely that their AI tools have resulted in unethical decisions since their implementation (Avanade, 2020). There was also strong consensus among the respondents (88%) as to the need for the regulation of ethical AI. Even in a country with a strong focus on data ethics, Denmark, survey data show that 78% of Danish small and medium-sized enterprises have not developed a policy or guidelines on data ethics (Danish Government, 2022). And yet there are benefits of implementing and deploying AI solutions in a responsible and ethical manner. Existing impact assessments of using ethical AI as opposed to unhinged AI show that 'an ethical framework would provide a net benefit, both from an economic perspective as well as for some of the 'softer' added value, such as projecting European values globally' (EPRS, 2020).

Accenture research found that firms categorised as 'AI achievers' – defined as those 'that have advanced their AI maturity enough to achieve superior growth and business transformation' – are more likely to take a responsible approach to data and AI than other firms with lower AI capabilities (Accenture, 2022). Adding to the picture, in a 2020 survey conducted by the Economist Intelligence Unit among 257 senior executives and IT decision-makers in the US, most

17 This includes respondents who reported that at least 20% of their organisation's earnings before interest and taxes in 2020 was attributable to AI use.

respondents recognised the importance of ethical AI for both talent acquisition and retention (The Economist Intelligence Unit, 2020). Other data sources indicate that an ethical corporate culture matters for workers and affects companies' ability to attract, recruit and retain employees (LRN, 2021; BBC, 2022). Recent survey data also suggest that ethics are particularly important for the younger generation (those born from 1997 onwards and generally labelled as 'Generation Z') when thinking about their future work environment (Tallo, 2021). This becomes a key consideration in the light of forecasts predicting that Generation Z will make up almost 30% of the total workforce in 2025 (WEF, 2022).

Referring to the US economy, economist Daron Acemoglu argues that 'ethical automation is not what we have been doing' (see MIT Sloan School of Management, 2021). Previous research suggests that the prevailing business model adopted by large tech corporations is mostly focused on cost-cutting and has led to over-automation (Acemoglu, 2021). This trend may have deprived society of many productivity gains that could have otherwise been achieved through automation and has been a primary driver of wage inequality in the US over the last four decades (Acemoglu and Restrepo, 2022). However, this logic can be reversed because the impact of technologies is not deterministic but results from deliberate choices; according to Acemoglu, 'AI could also be used for increasing worker productivity and expanding the set of tasks in which humans have a comparative advantage, rather than focusing mainly on automation' (see Boston Review, 2021). This suggests that government policies should be geared towards supporting AI that enhances jobs, rather than replacing or eliminating them.

While digital ethics is not yet a priority for many organisations, the 2022 Gartner Hype Cycle for AI suggests that this is likely to change, with 'responsible AI' expected to reach mainstream adoption within 5 to 10 years (Gartner, 2022). This will, however, only happen if there is a concerted effort from all stakeholders to leverage technologies for good purpose, in a way that creates opportunities for good-quality jobs.

Evidence from case studies

The main drivers of the adoption of digital technologies in all establishments interviewed for this report were process optimisation, efficiency and productivity gains, often in the context of competitive market pressure and persistent labour shortages. Other motives mentioned in the interviews included improvements in working conditions, for example to ensure that work was more rewarding, safer and less physically demanding for workers. Both the motives behind the adoption of the technology and the type of technology adopted shaped the approach to technological change taken in each establishment and the extent to which ethical issues were considered in the process of change.

In the case of the highly automated Hamburg-based Container Terminal Altenwerder (CTA), an additional and decisive factor that influenced the way the technology was adopted and implemented was the presence and involvement of an active works council and a strong union (ver.di), representing about 85% of the workforce. In all other establishments, employee involvement through formal employee representation structures (works councils or trade unions) was less apparent, as the introduction of the technologies was generally viewed positively by employee representatives because it reduced the need for workers to perform repetitive or physically demanding tasks and made work more rewarding for the employees concerned. Even in high-tech companies, automation technologies are seen not as threats to jobs but rather as supporting human activities and decision-making. However, both management and employee representatives acknowledged that further automation would require the greater involvement of worker representation due to the more significant changes to work organisation and working conditions that it might entail.

In most of the establishments interviewed, direct employee involvement was essential in the phase of technology design and implementation, as employees have an in-depth knowledge of the processes to be automated; this knowledge, accrued through practice, had to be incorporated into the set of instructions underlying the functioning of the automated systems. However, in some instances, the appropriation of workers' tacit knowledge gave rise to feelings of loss of pride or value in work or a sense of being somehow dispensable. From a management perspective, such concerns were addressed by intensifying communication with employees and being transparent about the purpose of the technology being deployed.

Apart from compliance with safety regulations and GDPR requirements, no explicit ethical policies, procedures or guidelines were used at any stage of technological change in any of the establishments interviewed. Ethical issues arising from technology adoption were to varying extents addressed as part of the implementation of GDPR protocols or broad digitalisation strategies. A more formal approach at group level is exemplified by the Data and AI Ethics Council, established in March 2021 in Orange. The council comprises 11 independent experts who are responsible for drawing up ethics guidelines for the responsible use of data and AI in the company and in the later implementation of these guidelines (Orange, 2021). These guidelines were, however, not explicitly mentioned in the Madrid-based establishment where the interviews were conducted.

Although training was provided to enhance employees' digital skills, it was largely intended to equip workers with the skills required to use or interact safely with the

specific technology adopted. There was limited evidence of comprehensive change management programmes implemented across the establishments interviewed. These can help workers to adapt better to new roles and address potential psychological impacts of technology adoption, including fear of future job losses, or anxieties around skills obsolescence, loss of expertise or reduced job crafting.[18]

The approach to technological change in all establishments was incremental, building on previous automation and digitisation efforts. Based on the interviews with management representatives, future plans invariably involve an increase in automation, and this is likely, in some cases, to entail some net job losses down the line. It is therefore essential for these establishments (and their workforce) to have a more forward-looking and ethical approach to digitalisation geared to augmenting human capacities rather than replacing jobs.

Advanced robotic systems and cobots

As automation was largely driven by the need to retain staff in an increasingly tight labour market, Dutch supplier of injection moulding products Wicro Plastics made a choice to automate only simple and repetitive manual tasks, leaving quality control tasks – including product control and approval – to human operators. The management assumed that this shift in tasks is more rewarding for employees, as they have more autonomy and their pace of work is less determined by machines. At the same time, the quality of products has improved, as the automated machines are more accurate in executing simple and repetitive tasks. Due to the proximity of human operators to machines on the shopfloor, the emphasis in technology adoption and implementation was largely on safeguarding workers' physical safety through a range of measures. These included thorough testing measures and risk assessments, routine maintenance, the installation of safeguarding devices, and the provision of safety training and guidelines.

Ergonomics and safety issues were also key considerations for the management of a Finnish medical device factory when deciding on which tasks to automate. One important objective of technology design and adoption was to relieve workers of repetitive and potentially hazardous tasks and make the work environment safer. The establishment has a prototype department, where all new technologies are tested before being implemented on the factory floor. There are several procedures, both formal and informal, to ensure that robots are introduced with safety in mind. All new technology needs to undergo a management of change procedure, whereby both the physical and the psychosocial work environments are assessed. All production lines undergo monthly risk assessment procedures called '6S-rounds', involving the environment, health and safety department, line managers, and occupational safety and health representatives.

In the Finnish establishment, the initial introduction of cobots proved to be a learning experience and was much less straightforward than expected. For example, to meet the safety requirements and standards of the production line, the first cobot introduced to the shopfloor was eventually caged and isolated from workers. More generally, according to the people interviewed in the establishment, the introduction of new automation technologies can be frustrating, as adjustments are often needed for the technology to fulfil its intended function. This also implies that piloting and risk assessment are processes that cannot be rushed and require time.

Strict compliance with safety standards and regulations was also a priority at Hamburg-based CTA in the context of increasing automation. Before newly developed systems or machinery are put into operation, these systems need to be certified and evaluated, which includes running a technological risk analysis as well as risk assessments. These tests are performed in accordance with the German Works Constitution Act, which guarantees the works council's co-determination rights in this respect, and in cooperation with the employer's liability insurance association (Berufsgenossenschaft). Beyond occupational safety and health, the priorities of the ver.di trade union and the works council were offering employment protection to the workforce and preserving good working conditions. The collective agreement signed in 2014 by the union and the management sets out a framework for dealing with technological change together in a socially responsible way. With the most recent agreement, concluded in 2022, the works council agreed to further automation, while the management committed to implementing employment protection measures and safeguarding employment until 2035. As part of this collective agreement, a new automation commission (consisting of employee and employer representatives) is tasked with developing a range of measures mainly related to requalification and training, work reallocation, working time reduction, and health and safety.

18 Job crafting refers to behaviour initiated by employees to shape or change the characteristics of a job to suit their personal needs, goals or skills.

In the other two establishments deploying advanced robots (and cobots), the digital transformation was neither monitored nor negotiated by employee representatives. Employee involvement instead took the form of direct participation in the prototyping and piloting of the robots/cobots; this was considered a necessity, as employees are the end-users of robotic applications. In the Finnish factory, employees were also involved in the 3D modelling and simulation of the technology, and the formulation and drafting of operating manuals. As robots are mostly designed and developed internally by automation engineers in Wicro Plastics, there were also several opportunities to involve employees at an early stage of the process of change. In both establishments, this employee involvement fostered, at least to some extent, greater acceptance of and trust in the technology adopted.

Training was another important component of the approach to technological change. While training in Wicro Plastics was provided primarily to ensure that human operators engaged safely with the automated machines, a more comprehensive approach was followed in the Finnish factory and the German container terminal. At the Finnish factory, additional training was offered to workers to equip them with the skills required to perform some basic maintenance work. The training offered to the workforce at CTA was more extensive, and is continuously adapted to newly emerging or changing job profiles. CTA also cooperates with the Maritime Competence Centre, which offers training and qualifications for seafarers and in the area of logistics, and access to programmes and funds of the German Federal Employment Agency (Bundesagentur für Arbeit).

Robotic process automation and artificial intelligence applications

Although technology adoption at the French infrared tech provider Lynred and telecommunication company Orange España was driven by the need to stay competitive on the market, the technology solutions were regarded as key to attracting and retaining skilled employees. The technologies deployed in Lynred and Orange España were AI and RPA applications, respectively, automating repetitive, tedious and time-consuming tasks involving a low level of decision-making and hence freeing up time for more rewarding and added-value tasks for the employees concerned.

The participative approach to technology adoption and deployment in these two establishments ensured a high level of employee buy-in and trust in the technological change. In Lynred, experimentation with new technologies is guided by suggestions from employees themselves.[19] Established in 2019, the Committee of Project 4.0 – composed of managers from different departments in two locations (Veurey-Voroize and Palaiseau) – reviews and selects digital projects proposed by employees based on a cost–benefit analysis and assessment of impacts on working conditions. The impact of the technology solution is also monitored and assessed on a regular basis after the technology is rolled out. This approach is routine and was also followed for the recent adoption and implementation of AI applications in the company's premises in Veurey-Voroize, including those facilitating the automated dispatching of helpdesk request tickets to the IT department and the automated calibration of a machine used in the unit producing bolometers.

A similar bottom-up approach to technology adoption was taken at the Madrid factory of Orange España, where employees can put forward their proposals to a robot factory for work processes that could be automated in their department. The Robot Factory endorses and oversees the development of RPA applications within the establishment. RPA applications are deployed widely in the establishment across different departments and for different functions, from HR to customer care and finance. Prior to any technology roll-out, the Robot Factory carries out a risk assessment, mainly focusing on data privacy and security, followed by a piloting process to check that the technology functions as intended. The robotic applications – when applied to non-critical work processes[20] – can also be developed by employees themselves if they complete robot 'self-management' training. After the employees complete the course, they receive a certificate as well as a licence to develop RPA solutions. Although this training is entirely optional, management promotes and encourages employees to complete it, as it is useful for their career prospects and helps each department to improve the efficiency of their processes. Initially, the training was conceived as an initiative for managing change, designed to foster greater acceptance of and trust in technologies among the workforce and, not least, empower employees to develop robotic applications and take ownership over the development process. In addition to the training, the Robot Factory coaches the employees developing RPA applications throughout the whole process, from design to deployment. All robotic applications are audited by the Robot Factory on a monthly basis to prevent any misuse and the malfunctioning of the technology.

19 In recognition of its efforts to leverage technologies to support employees' decision-making, in 2020 the company was awarded the label 'Industry of the Future Showcase' (Vitrine Industrie du Futur) by the Alliance Industrie du Futur (2020).

20 Critical processes are determined during the design phase through a risk assessment and a cost–benefit analysis, which discerns whether a risk is tolerable or not. A process is not defined as critical using standard criteria; its definition as such depends on the importance that the manager of the relevant area of the company gives to the process that is going to be automated, the data to be handled and the importance of the process for the company overall.

A top-down and more management-driven approach to technology adoption was instead taken by the Swedish municipality of Södertälje and the Danish unemployment fund Civiløkonomernes A-kasse (CA). In both cases, the approach was formalised in an overarching digitalisation strategy aimed at improving the accessibility of services and improving efficiency. The automation of work processes was also intended to reduce time spent by employees on repetitive administrative tasks, including data entry, processing, screening and retrieval, and better align tasks with employees' education and qualifications.

In the Södertälje municipality, a digitalisation unit conducts strategic oversight over the digitalisation processes across all municipal units and makes sure that digitalisation efforts in each unit are aligned with the overall municipal digitalisation strategy and integrated into its operational plan. Although the impetus for the deployment of both RPA technology in the payroll office and the AI chatbot in the contact centre came from the management, the technology solutions were designed, piloted and implemented in close collaboration with employees in the units concerned. The RPA technology is currently being applied to an increasing number of work processes and not only limited to the payroll functions. Involving employees in technology design and its fine-tuning improves their trust in the process and gives them a sense of control, as they know what the RPA technology does and how it works.

At the CA, the deployment of a machine learning-based system providing feedback on CVs and job applications (called Instant Feedback) kick-started the implementation of a new customer-centric digital strategy aimed at enhancing the availability of services to members. The preliminary screening of job applications, including the provision of routine feedback, lent itself to being automated, as it was a time-consuming and resource-intensive task for careers advisors and members increasingly required rapid feedback, often outside office hours. The implementation of the technological solution relied on the input of careers advisors on best practices for providing feedback. These were then translated into a set of standardised and codifiable rules and incorporated into the machine learning application, which was thoroughly piloted prior to its roll-out.

Ethics were not considered in their own right in the digitalisation strategies of CA and Orange España. However, training was provided to employees to enhance their digital skills, and continuous and targeted communication with employees was instrumental in addressing concerns of an ethical nature. These, for example, revolved around future job losses and accountability for inaccurate outcomes of automated systems.

Impact of technology in the workplace through an ethical lens

Evidence from survey data

Insight on areas of concern for employees about the future of work can be gleaned from a 2021 survey on ethics at work of 10,000 employees in 13 countries (including seven EU Member States) (IBE, 2021). The most reported concerns in the survey were the loss of interpersonal interactions due to the COVID-19 lockdown, followed by the misuse of AI for unethical behaviour, discrimination or biases in the workplace and the replacement of humans with automated machines or AI (Table 10). Of the EU Member States surveyed, employees in Portugal and Spain were the most likely to be concerned about most of the issues proposed in the survey, while employees in Germany and the Netherlands were least likely.

Table 10: Share of employees reporting concerns with regard to the future of their workplace in the seven surveyed EU Member States (%)

Concern	Spain	France	Germany	Ireland	Italy	Portugal	Netherlands	EU*
Loss of interpersonal interactions due to the COVID-19 lockdown	54	39	29	49	46	61	30	44
Misuse of AI for unethical behaviour (e.g. discrimination and privacy violations)	57	36	21	38	42	64	24	41
Discrimination or bias in the workplace	59	37	20	38	42	64	21	41
Automated machines or AI replacing humans in the workplace	59	40	22	39	40	61	26	40
Loss of interpersonal interactions due to new technologies	55	37	24	40	37	54	26	39
Inability of organisations to live up to their stated ethical standards	51	32	18	35	37	58	19	36
Increase in unethical behaviour due to an increase in the use of new technologies	53	35	17	33	35	57	18	35
Increased surveillance and monitoring in the workplace	41	35	28	41	29	42	22	34
New workplace/skillset requirements due to digitalisation and new technologies	47	36	19	37	30	39	22	33

Notes: *Figures correspond to the percentage of respondents who said that they were extremely concerned, moderately concerned or slightly concerned. * Unweighted average of the seven EU Member States covered in this survey.*
Source: *Second Delphi study, 2022*

Ethical concerns about AI use in the workplace were also explored in a survey conducted online by BCG GAMMA and Ipsos in 2018 among 7,077 people in the active population (18 years old and over) in seven countries (Canada, China, France, Germany, Spain, the United Kingdom and the US). Respondents to the survey generally saw AI as having a positive impact on organisations and themselves, for example by reducing the time employees spend on performing tedious tasks (72%), helping them meet deadlines and work faster (68%), and reducing the risk of errors (67%). However, respondents in France and Germany were less enthusiastic about the use of AI in the workplace.

As shown in Table 11, concerns that AI could result in an increase in supervisory control and surveillance in the workplace were higher in Spain and Germany than on average across all countries. Of the three EU Member States, respondents in France reported the highest levels of concern about job losses due to workers' replacement by machines, the dehumanisation of work and ethical issues related to data protection.

Table 11: Share of employees with concerns about specific workplace dangers of AI (%)

Concern	Spain	France	Germany	EU*	Average**
Result in more control and surveillance	81	73	79	78	76
Lead to job losses due to a reduced workload	67	69	66	67	68
Dehumanise work, resulting in less social cohesion	70	71	68	70	65
Pose ethical problems with regard to the protection of personal data	63	69	58	63	64

Notes: *Figures correspond to the percentage of respondents who answered 'yes' for the items listed in the question 'Do you think that, in your workplace, there is a danger that the development of AI and its applications may….' The total number of people interviewed was 1,018 in France, 1,010 in Germany and 1,009 in Spain. * Unweighted average of the three EU Member States covered in the survey. ** Average across all the countries included in the survey.*
Source: *BCG GAMMA and Ipsos, 2018*

A European survey conducted in 2018 among 1,400 managers found that digital privacy is the most pressing ethical concern arising from the use of digital technologies at work, and was reported by nearly half of respondents. Other ethical issues – for example, transparency, human dignity, accountability and human agency, discrimination and biases – were mentioned by respondents less often (CEC European Managers, 2018). Drawing on these findings, the pan-European organisation representing managerial staff in Europe (CEC European Managers) concluded that 'raising awareness on the risks and ethical implications is fundamental in times where orientation is scarce and where changes occur rapidly' (CEC European Managers, 2018, p. 11). The survey also highlighted challenges linked to working in a digital environment, as reported by managers. These were related to diminished work–life balance, and feelings of information overload, stress and not being up to date with the digital world.

When it comes to trust in AI applications at work among the general population, a global survey among 17,193 respondents aged 18 and over[21] conducted by the University of Queensland and KPMG Australia found that 48% of respondents are willing to trust AI at work but the level of trust varied between the countries surveyed (Gillespie et al, 2023). A lower-than-average level of trust in AI at work was, for example, reported in the EU Member States surveyed, namely Estonia (35%), Finland (26%), France (39%), Germany (37%) and the Netherlands (37%). Conversely, trust in AI is highest in emerging economies (66–87%).

Evidence from case studies

Advanced robotic systems and cobots

In the three establishments deploying advanced robotics solutions, the technologies had an impact on task content and largely resulted in the redistribution of tasks among operational staff. For example, in Wicro Plastics, operators on the shopfloor are spending an increasing amount of time on quality control tasks rather than assembly work. In addition, in a Finnish medical device factory, the robotic applications used in product assembly automate manual repetitive tasks that require a high level of precision or are unergonomic or hazardous for humans to do. However, in this establishment, the advanced robotic technologies require some basic maintenance that is performed by shopfloor operators, such as changing the glue that they use and the work surface on which they operate. The need for the supervision of machines and maintenance requirements have created a perception among shopfloor operators of robotic applications being labour inducing.

In addition, in the Hamburg-based CTA, automation has reduced the number of manual activities that workers need to perform – for example, operating gantry cranes or checking that containers are securely fastened. Operators have now been moved to a control room where they perform supervisory and controlling tasks. According to interviews, process analysis, optimisation and simplification through digital tools has helped to reduce stress and eliminated physically demanding tasks for terminal operators.

Contrary to expectations, technology adoption did not increase work intensity in any of the three establishments. At the Finnish factory, the cobot technology has instead slowed down the pace of work, as cobots perform work more accurately but at a slower pace than humans; this has reportedly created some frustration among employees.

The level of work autonomy of shopfloor or terminal operators is limited to start with; if anything, automation technologies have further reinforced the standardisation of work for operators working on assembly or production lines or at terminals. In the Finnish establishment, the deployment of advanced technologies has also increased the reliance of shopfloor operators on automation engineers or technicians to fix technical issues, which is perceived as reducing the autonomy of operators. This has also contributed to a feeling of disempowerment among operators when technical problems occur, as solutions are of a technical nature and are therefore beyond their competence. By contrast, in Wicro Plastics, with the tasks of shopfloor operators shifting to those associated with quality control, they perceive individual decision-making as having increased, as they more often need to assess whether a product meets quality standards. This has also created more interactions between operators and their line managers with regard to product quality control.

Some uneasiness about working with advanced collaborative systems was noted in the Finnish establishment. The chosen approach was therefore to deploy technologies and change workers' roles on a voluntary basis, in the sense that the advanced robot technologies are operated by employees who have volunteered to do so. Some employees did not welcome their work being reduced to robot supervision and maintenance, while others prefer performing such tasks over assembling products themselves.

Concerns around job replacement by machines in future was voiced to varying extents in the three establishments adopting advanced robotics applications. At the CTA, automation is regarded by

21 The survey covered 17 countries, including 5 EU Member States, and the sample size across countries ranged from 1,001 to 1,021 respondents.

works council representatives as a threat to employment. Employees are in favour of automation only if there is a degree of employment security, which, in the establishment, is provided by an agreement on protection from dismissal until 2035. The fear of future job losses due to automation was less apparent in Wicro Plastic, where robots were deployed partly to compensate for staff shortages.

Other concerns of an ethical nature that surfaced in the investigation of the case studies revolved around loss of job crafting and pride in work previously performed by operators. For example, in the Finnish establishment, some employees raised concerns about whether a cobot could create as good a product as they could themselves with years of experience in manufacturing.

Finally, in the establishments interviewed, there was some recognition that greater automation would have an impact on the quality of work to a much larger extent than seen so far. In this regard, the works council representative interviewed at CTA noted that the full automation of certain activities is expected to involve significant changes to the nature of operators' work and is a development that many employees might not be comfortable with.

Robotic process automation and artificial intelligence applications

In the establishments interviewed, the adoption of RPA and AI technologies has improved working conditions and, in spite of some initial fears, has so far not led to job losses. If anything, automation resulted in the creation of some new technical jobs. In all four establishments deploying RPA and AI solutions, the technologies have diminished the need for employees to perform time-consuming and repetitive or manual tasks and freed up time for them to perform more added-value and rewarding tasks. In some cases, as in Lynred and Orange España, the adoption of technology has also made it possible to insource some processes that had been previously subcontracted to external providers.

In Lynred, the two AI-based applications have not only automated routine tasks but also resulted in the reallocation or redefinition of some tasks. For example, the AI solution has automated the calibration of parameters of a machine used in the bolometer production unit, which was previously done manually by the supporting engineering team. These parameters can be used directly by production operators without relying on the supporting engineering team, thus reducing downtime and streamlining the production process. In the case of the automation of IT helpdesk ticket dispatching, the first-level support team were relieved of the task of manually dispatching requests to other more specialised IT support teams, and the time saved is reallocated to other tasks such as the set-up of IT equipment, including computers. According to management, the manual dispatching of tickets is a repetitive and not very fulfilling job, which was one reason for the previously high turnover within the team.

The RPA technology at the payroll office in the Södertälje municipality also saves time for employees and assists them during periods when workload is at its peak – that is, when salaries are paid out. In this department, the technology has had an impact on job content and prompted a shift from manual administrative work to coordination and supervisory tasks. While this change was welcomed by some office employees, others consider that the supervision of robotic processes and quality control are not part of their job description. However, loss of human agency over work processes was not an issue, as RPA is used to support human decision-making rather than automate decisions.

The AI chatbot in the municipality contact centre was well received by all employees and regarded as a rewarding and novel tool to work with. Contact centre employees have access to a tool enabling them to revise or fine-tune the chatbot answers if they see from logs that it does not understand questions correctly or its answers are not accurate. Nevertheless, the amount of contact they have with the public through phone calls has not changed; according to municipal statistics, the chatbot is mainly used by citizens outside opening hours, in the evenings and at weekends. As the chatbot is connected to an IoT network, it also provides information that would be difficult for human employees to provide, such as real-time data on water temperatures at bathing sites and air quality, and information about free spaces in car parks.

In all establishments, most notably Orange España, technology adoption goes hand in hand with continuous training and skills development. As part of a collective agreement, the company is committed to providing continuous training to all employees to upgrade their skills and enable them to adapt to technological change (Ministerio de Trabajo, Migraciones y Seguridad Social, 2019). Orange España also has a training committee – where trade union and works council representatives are also members – that performs a periodic evaluation of the training programmes and agrees on the training to be provided in the best interest of both the company and the workers. In the case of the RPA technology, training in robot self-management fostered greater acceptance of the technology among the workforce and was instrumental in addressing employees' concerns around data protection and security, as well as accountability, and their ownership over RPA processes.

In other establishments, both training provision and employees' involvement in the design, development and piloting of technology applications not only enhanced their digital competence and skills, but also improved their understanding of the functioning of the

technology, including its limitations. This helped, for example, to dispel fears of being replaced by machines, which were expressed by employees in most establishments in the initial phase of the technology introduction. This better understanding of the technology solutions contributed to addressing other concerns as well. For example, the introduction of the Instant Feedback system at the CA was initially regarded by careers advisors as a tool that would diminish the value of their work. Careers advisors were also concerned about unclear accountability for the quality of the system's output. As they became more involved in the design and development of the application, it became apparent that the technology solution would be focused on providing standardised and routine advice regarding CVs and job applications and matching users with job advertisements. The deployment of the Instant Feedback system has changed the job content of careers advisers for the better, enabling them to spend more time on providing tailored advice, support and coaching to jobseekers.

Another concern raised by those at Orange España and the Södertälje municipality in relation to the RPA technology related to occasional glitches or malfunctions that may be stress-inducing, especially in situations where employees feel that they have lost expertise or lost control over tasks that have been automated for some time and that they no longer perform manually. This is one of the reasons why the Robot Factory in Orange España does not allow the complete automation of any process and ensures that employees in charge of an automated process retain some expertise to perform the task manually.

With regard to privacy and data protection, all the establishments interviewed follow strict protocols to comply with the GDPR at all stages, from the design of the product or service to its implementation. While the automated processes generate vast amounts of data, these are anonymised and are not used for monitoring employees' performance or surveillance purposes. In Orange España, the principle of 'privacy by design' is routinely applied in the design and development of any technology. Other ethical principles – such as human oversight and 'human in the loop' – are also taken into account in technology design in other establishments and their importance is reiterated in the communication with employees around the adoption of the technology and emphasised in the provision of training.

Key takeaways

- Except for data protection policies and health and safety procedures ensuring compliance with the law, there are no explicit policies and procedures for addressing a wider range of ethical issues arising from technology use in the establishments interviewed. This finding resonates with findings from various surveys on AI use in the workplace pointing to a lack of ethical policies guiding technology implementation and use. In the selected establishments, ethical principles were generally not considered in their own right but were loosely embedded in technical policies and procedures or broader digitalisation strategies. From a policy perspective, further guidance and incentives for companies may be needed to foster the ethical and responsible design and use of technologies in the workplace. There may also be constraints on ethical technology design, development and use, and further investigation is required to better understand them.
- Ethical concerns raised and addressed in the selected establishments during technology adoption and implementation are often related to employees' fear of being replaced by machines in the non-distant future, questions around accountability and transparency in automated processes, and trust in human–machine collaborative systems. In the establishments interviewed that use machine learning, the applications raised fewer concerns, as these involved the automation of routine and codifiable tasks requiring low-level decision-making. The participative approach to technology adoption taken in these establishments also helped to defuse any ethical concerns outright.
- Trust in the transparency of and the reliability of technology were core concerns for management in both technology design and use. Key GDPR principles – such as transparency, purpose, data limitation and security – remain key to an ethical and trustworthy approach to technology design and use in the workplace. However, ever more powerful and sophisticated AI technologies may challenge some GDPR principles, such as transparency and data minimisation, as machine learning systems are complex and feed off large amounts of data (more than may be necessary).
- The approach taken to automation in establishments tends to be incremental, building on previous digitalisation efforts and aimed at increasing automation. Organisations that have already adopted automation and AI technologies should determine which work processes are most likely to be automated in the future and put in place forward-looking strategies to support employees expected to be most affected by technological change.

4 Conclusions and policy pointers

Conclusions

Technologies are becoming increasingly smarter, and are edging their way into the workplace. The discourse around AI and other smart technologies has so far been rather ambivalent. While it is recognised that new digital technologies have the potential to improve productivity, efficiency and accuracy, there are mounting concerns about negative side-effects of new and emerging technologies for work and employment if they are not designed, developed and used in the workplace ethically. Ethical issues arising from new technologies and their negative impacts on working conditions often originate from deliberate choices made around technology design and implementation. Other important factors determining the impact of digital technologies in the workplace are the extent to which workers and their representatives are involved in the process of change and policy responses to the challenges posed by digitalisation.

Ethical concerns often relate to potential data protection and privacy breaches, particularly in the context of intrusive data-driven work management and employee monitoring practices. Among the most acknowledged risks of using AI tools – for instance for recruitment and performance management – is the risk of biases deepening workplace inequalities and opening the door to discriminatory practices. Other ethical issues brought up by experts and stakeholders consulted in relation to the use of AI systems revolve around accuracy, transparency, accountability and responsibility, particularly when technologies do not work as expected or their decision-making logic is difficult to trace or understand. These considerations are connected to concerns around explainability and the trustworthiness of AI technologies more generally. A lingering concern is also that technologies could sideline humans, potentially leaving them out of the loop, or, even worse, dehumanise work, compromise human dignity and eventually replace human workers altogether.

However, the automation of work and job replacement by machines are not new; they already took place in the previous industrial revolution, although they were confined to the physical domain. From a worker's perspective, the unfolding digital revolution may be more daunting, as automation and digitisation increasingly extend to the cognitive domain. Intelligent systems – with capabilities for self-learning and adaptation – can perform repetitive and physically demanding tasks but are also increasingly making inroads into functions requiring greater cognitive ability. Although the technology applications in most establishments interviewed are limited to the automation of routine and codifiable tasks requiring low-level decision-making and have not (yet) eliminated jobs, employees often voiced concerns about future job losses or losing knowledge or skills. There are also concerns about the ownership and appropriation of knowledge, with technologies leveraging workers' tacit knowledge, which ultimately comes from their accrued practice and work experience. This same knowledge gives workers a sense of identity and belonging in the workplace and pride in their work.

With the pace of work automation and digitisation increasing, changes to ways of working and the content of work will mean that some jobs may no longer be needed, while others may be replaced by new jobs requiring more advanced technical skills. With the more widespread adoption of AI in workplaces, some legal uncertainties may also emerge about protections granted to employees whose roles are automated or replaced by AI.

There are also fears that the use of new digital technologies may contribute to workers' deskilling, as seen in retail and business logistics, where workers are constantly monitored and directed by automated machines as to what tasks to do and how to perform them. This way of working also underpins the functioning of much platform work and is typically referred to as algorithmic work management. The ethical implications of such working arrangements are particularly apparent, as human agency and autonomy is significantly compromised and the meaningfulness of work diminished.

There is no doubt that increasingly sophisticated AI technologies will have far-reaching implications for all classes of workers, including well-trained and experienced professionals. A case in point is generative AI, such as ChatGPT, which has the potential to transform many knowledge-intensive jobs. As suggested by recent forecasts, generative AI may not just displace jobs but also create new demand for labour. This may include less desirable jobs, for example dealing with the sheer volume of information the technology generates. With the ever-increasing availability of data, storage capacity and computing power, technologies are bound to improve and become more powerful. This is likely to amplify existing concerns around, for example, biases, inaccuracies, lack of transparency and potential misuse.

There is a need for greater awareness about the ethical implications of workplace technologies and for policies promoting an ethical and human-centric approach to technology design, development and use in the workplace. Some of the most disruptive effects of technologies on work can be offset if an ethical and human-centric approach is incorporated into technology design, development and use in the workplace. This presupposes, however, that there is a will to ensure that technologies are leveraged 'for good purpose', to enrich working life, augment human capacities, create new roles and support human values, instead of the focus being exclusively on cost efficiency.

Whether developing or buying a technology solution, it is critical to take an iterative approach whereby technologies are regularly assessed through an ethical lens, particularly when technologies are repurposed for uses other than those initially intended and communicated to the workforce. Ethical principles should also be reviewed on a regular basis, as technologies evolve quickly and become more powerful. This approach to technology design and implementation should be accompanied by sound change management practices that foster workers' resilience and adaptation in a technology-driven and constantly changing work environment. Ethics training should be provided as part of change management, to encourage a common understanding of agreed ethical principles within organisations and promote awareness among the workforce of the ethical implications of technologies.

As highlighted by the experts consulted and as suggested by survey and case study evidence, approaches to workplace digitalisation do not systematically incorporate ethical thinking, showing some disconnect between the discourse among experts on technology ethics and business practices on the ground. The digital transformation of work is a process capable of generating irreversible changes to several aspects of working life; it is how technologies are implemented in the workplace that makes the difference in ensuring that both businesses and workers can truly benefit from them.

Trust should function as a cornerstone of ethical practice in technology adoption and use. Trust emerged as an important concern when it comes to the digitalisation of work, both from the consultation with stakeholders and from the interviews with establishments. Ethical practice in technology adoption and use is key to promoting trust at all levels of governance and fostering general acceptance of technologies. Public authorities have a role to play in this; they must deploy technologies ethically in their own practices and ultimately be the gatekeepers of the ethical use of technologies in society. They should start by committing to avoiding the ethical pitfalls that have been seen in controversial cases of algorithmic systems used in public service delivery. Such cases have received a great deal of negative media attention, fuelling widespread concerns about the transparency and accountability of AI systems and the arbitrariness of algorithmic decisions. Ethical principles and democratic values should be upheld at all times in governments' digital practices as much as in businesses' digital practices (European Parliament, 2022b).

At establishment level, employee involvement is a crucial requirement for building workers' trust in technologies and facilitating their adaptation to new ways of working. Ethical guidelines issued by social partners routinely flag employee involvement and participation as a prerequisite for the successful digital transformation of the workplace in the dual interest of promoting workers' well-being and enhancing productivity and efficiency. Using co-creation models could therefore be a promising way of designing and introducing new technologies in the workplace.

Social dialogue and collective bargaining are central to an ethical, responsible and fair approach to technology adoption in the workplace. Based on the interviews with stakeholders, there is convergence in the interests of social partners, particularly when it comes to enhancing the provision of training and upskilling for the workforce to enable them to adjust to technological change. However, social dialogue has a bigger role to play, beyond skills issues; greater efforts are required from social partners to address new ethical challenges (with ramifications for quality of work) in a fast-paced digital world of work.

Research shows that collective bargaining is somewhat slow to adapt to complex new realities involving digital innovation and the disruption of work. Innovations in collective bargaining are predominantly incremental, updating traditional industrial relations topics and processes in response to changing conditions. There are nonetheless instances of innovative sectoral and company-level collective agreements introducing new digital rights or safeguards to address the power imbalance in the workplace that ubiquitous or intrusive technologies can create or contribute to deepening. Collective agreements at company and sector levels are particularly effective and useful tools to address the most pressing concerns arising from the automation and digitisation of work.

Policy pointers

- Initiatives to promote ethical thinking and practices in technology design and use in the workplace should be added to the catalogue of measures included in national digital and AI strategies. These might include campaigns to raise awareness of the benefits of ethical workplace digitalisation or guidance for companies on embedding ethics in their digital practices and moving towards adopting a 'corporate digital responsibility' model. Such initiatives – accompanied by concrete investment plans – will help to bring the workplace dimension into greater focus in digitalisation policies and draw attention to important quality-of-work issues – for example, human agency, workplace fairness and equality – that are ethical issues at their core.
- It would be beneficial to incorporate ethics into broader digital education initiatives and policies aimed at supporting human capacity building in the area of AI and reforming the formal education system. Ethics should be embedded not only in standard curricula on computing and engineering but at all levels of the education system. This would raise the public's awareness of the potential ethical pitfalls of new technologies, improve their critical thinking about these technologies' potential ethical implications and cultivate their ethical sensitivity, as citizens, workers and developers of technology applications.
- Public–private partnerships can be instrumental in establishing ethical standards and governance frameworks for the responsible and human-centric development and use of technologies. Examples of such partnerships are 'regulatory sandboxes' for ethical AI. These have the potential to enhance ethical compliance, support knowledge sharing and foster responsible innovation. An example is the regulatory sandbox launched by the Spanish government (in collaboration with the European Commission) and open to other EU Member States. The initiative is aimed at developing practical guidelines and tools for the development and use of high-risk AI systems and exploring these systems' compliance with the upcoming AI Act. Other broad partnerships – involving the research community and often initiated by governments, as seen in Belgium, Germany and the Netherlands – should be further incentivised, as they contribute to the development of best practices and ethical standards and methods, and they create awareness of both the challenges and the opportunities that AI brings to our society (and the labour market). The involvement of social partners in such partnerships would ensure a balance of interests in the development of ethical approaches to AI in the context of work.
- Ethical digitalisation in the workplace should go beyond the mitigation of ethical issues as they arise during the implementation of technology. Rather, it should involve developing and deploying purposeful technology applications in line with core corporate values firmly grounded in compliance and ethics. This applies not exclusively to high-risk applications but more generally to digital projects. Ethical concerns and ways to address them should be considered as important as other corporate values, such as environmental, social and corporate governance principles and standards. Along these lines, the Danish government, a frontrunner in data ethics, has launched policy initiatives to encourage companies to use data (and technologies) in an ethical and responsible way in their business activities. Such an approach is in the interest of all stakeholders. A culture of trust and ethical business practices matters for employees and contributes to making jobs more meaningful. From an employer's perspective, a strong ethical corporate culture is essential for attracting and retaining employees, especially in times of persistent labour shortages, while enhancing brand reputation and building trust among shareholders and investors. As a step in this direction, organisations might consider appointing a chief ethics officer (or equivalent) tasked with promoting compliance with ethical standards in technology design and use.
- Initiatives taken by governments (for example, in Spain) and non-profit organisations (for example, the German AI Association) are experimenting with quality labels or trust seals for technology solutions (including low-risk applications), certifying that the technology is legitimate, ethical in its design and can be trusted. There is, however, a risk that these seals could result in ethics washing – creating the illusion that ethical issues are being addressed in technology design.
- Company-level collective agreements are among the most appropriate tools to address ethical issues arising from constantly evolving digital technologies. Social partners and negotiating parties in collective bargaining could benefit from enhanced exchange of good practices and information. Addressing ethical issues – such as around the explainability, transparency and accountability of AI systems – requires a coordinated and hands-on approach, for example involving the exchange of model agreements and clauses and the development of practical guides (beyond recommendations, ethical principles and checklists) enabling both management and employee representatives to deal with the challenges posed by the digitalisation of work.

- It may be beneficial for trade union and employee representatives to develop their technical expertise so that they are better prepared to identify and address new ethical challenges arising from the digitalisation of work. The field of AI is very complex, employing advanced concepts and techniques that are constantly evolving and often difficult to grasp for those with a non-technical background. Consideration might be given to the inclusion of provisions in collective agreements or legal frameworks – as exemplified in the German Works Council Modernisation Act – that give trade unions and employee representatives the opportunity to access external technical expertise to ensure that transparency and explainability are not compromised in the implementation and use of AI-powered technologies.

References

All Eurofound publications are available at https://www.eurofound.europa.eu

ABVV-FGTB (General Labour Federation of Belgium) (2019), *Ensemble pour une digitalisation socialement responsible/Samen voor sociaal verantwoorde digitalisering*, Brussels.

Accenture (2022), *The art of AI maturity: Advancing from practice to performance*, web page, available at https://www.accenture.com/gb-en/insights/artificial-intelligence/ai-maturity-and-transformation, accessed 20 June 2023.

Acemoglu, D. (2021), *Harms of AI*, NBER Working Paper No. 29247, National Bureau of Economic Research, Cambridge, MA.

Acemoglu, D. and Restrepo, P. (2022), 'Tasks, automation, and the rise in U.S. wage inequality', *Econometrica*, Vol. 90, No. 5, pp. 1973–2016.

Actu IA (2022), 'LaborIA: Matrice launches a survey on the impact of artificial intelligence on work this September', blog post, 2 August.

addAI (2021), *Sammanfattning: Hur mycket bör chefen veta om dig med hjälp av AI?* web page, available at https://www.addai.org/index.php/2021/03/28/sammanfattning-hur-mycket-bor-chefen-veta-om-dig-med-hjalp-av-ai/, accessed 20 June 2023.

AEPD (Spanish Data Protection Authority) (2020), *GDPR compliance of processings that embed artificial intelligence: An introduction*, Madrid.

AEPD (2021), *Audit requirements for personal data processing activities involving AI*, Madrid.

AFB, FEC FO, CFTC banques, SNB CFE-CGC, CFDT banques assurances (2020), Convention collective nationale de la banque du 10 janvier 2000. Etendue par arrêté du 17 novembre 2004 JORF 11 décembre 2004 – Textes attachés – Accord du 27 mai 2020 relatif à la gestion prévisionnelle des emplois et des compétences (GPEC).

AI HLEG (High-Level Expert Group on Artificial Intelligence) (2019), *Ethics guidelines for trustworthy AI*, Publications Office of the European Union, Luxembourg.

Akava, Confederation of Finnish industries, Commission for Church Employers, KT Local Government Employers, Central Organisation of Finnish Trade Unions SAK, Finnish Confederation of Salaried Employees STTK, Office for the Government as Employer VTML (2019), *Digitalising Finland is an opportunity: A big leap forward in employee wellbeing and in labour productivity*, Helsinki.

Alliance Industrie du Futur (2020), 'L'Alliance Industrie du Futur labellise 4 nouvelles entreprises "Vitrines Industrie du Futur"', press release, 22 June.

Avanade (2020), 'One in three Belgian C-suite executives admits their company's AI-tools have most likely made unethical decisions', press release, 22 July.

BBC (2022), 'Are workers really quitting over company values?' blog post, 28 February.

BCG GAMMA and Ipsos (2018), *Artificial intelligence: Have no fear – The revolution of AI at work*.

BDA (Confederation of German Employers' Associations) (2020), *Germany reloaded: Wie Wirtschaft und Beschäftigte von der Digitalisierung profitieren können*, Berlin.

Boston Review (2021), 'AI's future doesn't have to be dystopian', blog post, 20 May.

Brookings (2022), *The EU AI Act will have global impact, but a limited Brussels effect*, web page, accessed 4 July 2021.

Buiten, M. (2019), 'Towards intelligent regulation of artificial intelligence', *European Journal of Risk Regulation*, Vol. 10, No. 1, pp. 41–59.

Bynum, T. W. (2006), 'Flourishing ethics', *Ethics and Information Technology*, Vol. 8, pp. 157–173.

Capurro, R. (2017), 'Digitisation as an ethical challenge', *AI and Society*, Vol. 32, pp. 277–283.

CC.OO, Ametic and UGT (2020), *Recomendaciones sobre el Impacto de la Tecnología en los Centros Productivos de Trabajo*, Madrid.

CEC European Managers (2018), *Management in the digital era: Report on the survey results*, Brussels.

Ceemet and IndustriAll Europe (2020), *The impact of digitalisation on the world of work in the MET industries: An IndustriAll Europe and Ceemet joint position*, Brussels.

Ceemet and IndustriAll Europe (2023), *Artificial intelligence in the MET industries: IndustriAll Europe and Ceemet joint conclusions*, Brussels.

CEOE (Confederación Española de Organizaciones Empresariales) (2020), *Plan digital 2025: La digitalización de la sociedad española*, Madrid.

CGSLB (General Confederation of Liberal Trade Unions of Belgium) (2018), *Digitale transformatie: Kansen en bedreigingen. Studie over de impact van de digitalisering op werk*, Brussels.

CNPEN (National Pilot Committee for Digital Ethics) (2021), *Ethical issues of conversational agents*, Opinion No. 3, Paris.

Confederation of Swedish Enterprise (2018), *Artificial intelligence (AI) and its applications*, Stockholm.

Confederation of Swedish Enterprise (2020a), *Framtidens digitala arbetsmarknad*, Stockholm.

Confederation of Swedish Enterprise (2020b), *Vad innehåller parternas överenskommelse*, Stockholm.

Council of the EU (2022a), 'Artificial Intelligence Act: Council calls for promoting safe AI that respects fundamental rights', press release, 6 December.

Council of the EU (2022b), *Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts – General approach*, Brussels, 25 November 2022.

Danish Government (2022), *National strategy for digitalisation: Together in the digital development*, Ministry of Finance, Copenhagen.

DataDrivenInvestor (2018), 'Symbolic AI v/s non-symbolic AI, and everything in between?' blog post, 19 October.

D'Cruz, P., Du, S., Noronha, E., Parboteeah, K. P., Trittin-Ulbrich, H. and Whelan, G. (2022), 'Technology, megatrends and work: Thoughts on the future of business ethics', *Journal of Business Ethics*, Vol. 180, pp. 879–902.

Deloitte (2020), *Ethics and the future of work*, web page, accessed 4 July 2023.

Deloitte (2022), *State of ethics and trust in technology: Annual report*, 1st ed.

De Stefano, V. and Taes, S. (2023), 'Algorithmic management and collective bargaining', *Transfer*, Vol. 29, No. 1, pp. 21–36.

DGB (Deutscher Gewerkschaftsbund) (2020), *Artificial intelligence (AI) for good work*, Berlin.

DGB (2021), Stellungnahme des Deutschen Gewerkschaftsbundes zu dem Referentenentwurf des Bundesministeriums für Arbeit und Soziales (BMAS). Entwurf eines Gesetzes zur Förderung der Betriebsratswahlen und zur Stärkung der Betriebsräte (Betriebsrätestärkungsgesetz).

DGB (2022), 'Vertrauen ist gut, Rechte sind besser', press release, 9 February.

DI (Dansk Industri) and 3F (United Federation of Workers in Denmark) (2019), *Fremtidens transport, handel og logistik*, Copenhagen.

Diario La Ley (2023), 'Publicado el borrador del Real Decreto regulador del sandbox de inteligencia artificial', blog post, 2 June.

EDPB (European Data Protection Board) (2020), 'Hamburg Commissioner fines H&M 35.3 million euro for data protection violations in service centre', press release, 2 October.

EDPB (2022), 'The French SA fines Clearview AI EUR 20 million', press release, 20 October.

EDPB and EDPS (European Data Protection Supervisor) (2021), *Joint Opinion 5/2021 on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*, Brussels.

EDPS (2015), *Opinion 4/2015: Towards a new digital ethics – Data, dignity and technology*, Brussels.

EDPS (undated), *Data protection*, web page, available at https://edps.europa.eu/data-protection/data-protection_en, accessed 20 June 2023.

El País (2020), 'Yolanda Díaz: "El avance tecnológico no puede suponer un retroceso en las condiciones laborales al siglo XIX"', 25 July.

EPRS (European Parliamentary Research Service) (2020), *European framework on ethical aspects of artificial intelligence, robotics and related technologies: European added value assessment*, Brussels.

EPRS (2022a), *AI and digital tools in workplace management and evaluation: An assessment of the EU's legal framework*, Brussels.

EPRS (2022b), *Auditing the quality of datasets used in algorithmic decision-making systems*, Brussels.

ETNO (European Telecommunications Network Operators Association) and UNI Europa ICTS (2020), *The telecom social dialogue committee joint declaration on artificial intelligence*, Brussels.

ETUC (European Trade Union Confederation) (2022), *ETUC resolution calling for an EU directive on algorithmic systems at work*, Brussels.

ETUC, BusinessEurope, CEEP and SMEunited (2020), *European social partners framework agreement on digitalisation*, Brussels.

ETUC, BusinessEurope, SGI Europe and SMEunited (2021), *Implementation of the ETUC/BusinessEurope/SMEunited/SGI Europe framework agreement on digitalisation – 1st joint report*, Brussels.

ETUC, BusinessEurope, SGI Europe and SMEunited (2022), *Implementation of the ETUC/BusinessEurope/SMEunited/SGI Europe framework agreement on digitalisation – 2nd joint report*, Brussels.

Eurofound (2017), *Sixth European Working Conditions Survey – Overview report*, Publications Office of the European Union, Luxembourg.

Eurofound (2018a), *Automation, digitisation and platforms: Implications for work and employment*, Publications Office of the European Union, Luxembourg.

Eurofound (2018b), *Advanced industrial robotics: Taking human-robot collaboration to the next level*, Eurofound working paper, Dublin.

Eurofound (2019), *Advanced robotics: Implications of game-changing technologies in the services sector in Europe*, Eurofound working paper, Dublin.

Eurofound (2020a), *Employee monitoring and surveillance: The challenges of digitalisation*, Publications Office of the European Union, Luxembourg.

Eurofound (2020b), *Game-changing technologies: Transforming production and employment in Europe*, Publications Office of the European Union, Luxembourg.

Eurofound (2021a), *Digitisation in the workplace*, Publications Office of the European Union, Luxembourg.

Eurofound (2021b), *Impact of digitalisation on social dialogue and collective bargaining*, research digest, Dublin.

Eurofound (2021c), *Right to disconnect: Exploring company practices*, Publications Office of the European Union, Luxembourg.

Eurofound (2022a), *Ethics in the digital workplace*, Publications Office of the European Union, Luxembourg.

Eurofound (2022b), *Moving with the times: Emerging practices and provisions in collective bargaining*, Publications Office of the European Union, Luxembourg.

Eurofound (2023), *Right to disconnect: Implementation and impact at company level*, Publications Office of the European Union, Luxembourg.

European Commission (2020a), *White paper on artificial intelligence – A European approach to excellence and trust*, COM(2020)65 final, Brussels.

European Commission (2020b), *European enterprise survey on the use of technologies based on artificial intelligence*, Publications Office of the European Union, Luxembourg.

European Commission (2021a), *Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts*, COM(2021)206 final, Brussels.

European Commission (2021b), *Speech by Executive Vice-President Vestager at the press conference on fostering a European approach to artificial intelligence*, web page, available at https://ec.europa.eu/commission/presscorner/detail/en/speech_21_1866, accessed 20 June 2023.

European Commission (2021c), *EU Strategic Framework on Health and Safety at Work 2021–2027: Occupational safety and health in a changing world of work*, COM(2021)323 final, Brussels.

European Commission (2021d), *Proposal for a regulation of the European Parliament and of the Council on machinery products*, COM(2021)202 final, Brussels.

European Commission (2021e), *Inception impact assessment: Civil liability – Adapting liability rules to the digital age and artificial intelligence*, Brussels.

European Commission (2022), *Proposal for a directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)*, COM(2022)496 final, Brussels.

European Parliament (2021), *Resolution of 21 January 2021 with recommendations to the Commission on the right to disconnect*, 2019/2181(INL), Brussels.

European Parliament (2022a), *Resolution of 5 July 2022 on mental health in the digital world of work*, 2021/2098(INI), Strasbourg.

European Parliament (2022b), *Resolution of 3 May 2022 on artificial intelligence in a digital age*, 2020/2266(INI), Strasbourg.

European Parliament (2023), *Draft compromise amendments on the draft report Proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts*.

Eurostat (2023), *Use of artificial intelligence in enterprises*, web page, available at https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Use_of_artificial_intelligence_in_enterprises, accessed 20 June 2023.

Expert Group on Liability and New Technologies (2019), *Liability for artificial intelligence and other emerging technologies*, Publications Office of the European Union, Luxembourg.

Finnish Government (2018), *Eettistä tietopolitiikkaa tekoälyn aikakaudella: selonteko*, Helsinki.

Foffano, F., Scantamburlo, T. and Cortés, A. (2023), 'Investing in AI for social good: An analysis of European national strategies', *AI & Society*, Vol. 38, pp. 479–500.

Forbes (2019), 'What is the artificial intelligence of things? When AI meets IoT', blog post, 20 December.

FRA (European Union Agency for Fundamental Rights) (2018), *#BigData: Discrimination in data-supported decision making*, FRA Focus, Vienna.

FRA (2019), *Data quality and artificial intelligence – Mitigating bias and error to protect fundamental rights*, FRA Focus, Vienna.

FRA (2020), *Getting the future right – Artificial intelligence and fundamental rights*, Publications Office of the European Union, Luxembourg.

FRA (2022), *Bias in algorithms – Artificial intelligence and discrimination*, Publications Office of the European Union, Luxembourg.

Friedman, B., Kahn, P. H. and Borning, A. (2008), 'Value sensitive design and information systems', in Himma, K. E. and Tavani, H. T. (eds.), *The handbook of information and computer ethics*, John Wiley & Sons, Hoboken, New Jersey, pp. 69–101.

Friedrich-Ebert-Stiftung (2022), *Algorithmic management – Awareness, risks and response of the social partners*, Brussels.

Future of Life Institute (undated), *The AI Act – Analyses*, web page, available at https://artificialintelligenceact.eu, accessed 20 June 2023.

Futurion (2020), *Mattias Beijmo: Så övervakas vi av våra arbetsgivare*, web page, available at https://futurion.se/mattias-beijmo-sa-overvakas-vi-av-vara-arbetsgivare/, accessed 20 June 2023.

Gartner (2022), 'What is new in artificial intelligence from the 2022 Gartner Hype Cycle', blog post, 15 September.

Genesys (2019), 'New workplace survey finds nearly 80% of employers aren't worried about unethical use of AI – But maybe they should be', press release, 4 September.

Gillespie, N., Lockey, S., Curtis, C., Pool, J. and Akbari, A. (2023), *Trust in artificial intelligence: A global study*, University of Queensland and KPMG Australia.

Godino, A., Sander, J. and Molina, O. (2022), *Developments in algorithmic management from an IR-perspective: Spain*, Incoding Stock-taking Reports.

Goldman Sachs (2023), 'The potentially large effects of artificial intelligence on economic growth', blog post, 26 March.

Government of Spain (2021), 'The government adopts the Digital Rights Charter to articulate a reference framework to guarantee citizens' rights in the new digital age', press release, 14 July.

Government of Sweden (2021), *A good work environment for the future – The government's work environment strategy 2021–2025*, Government Communication 2020/21:92, Stockholm.

GPDP (Garante per la Protezione dei Dati Personali) (2021), Ordinanza ingiunzione nei confronti di Proma S.S.A. s.r.l., Injunction Order No. 136, 15 April.

Hansson, S. O. (2017), *The ethics of technology: Methods and approaches*, Rowman & Littlefield International, New York.

IAPP (International Association of Privacy Professionals) (2023), *Dutch DPA to enhance algorithm supervision*, web page, accessed 4 July 2023.

IBE (Institute of Business Ethics) (2021), *Ethics at work: 2021 international survey of employees*, London.

ILO (International Labour Organization) (2022), *ILO declaration on fundamental principles and rights at work*, Geneva.

Ingenjören (2020), 'Övervakning på jobbet ökar – också i Sverige', 23 September.

Inria (National Institute for Research in Digital Science and Technology) (2022), 'Artificial intelligence: What impact will it have on the world of work?' blog post, 27 September.

Jobin, A., Ienca, M. and Vayena, E. (2019), 'The global landscape of AI ethics guidelines', *Nature Machine Intelligence*, Vol. 1, No. 9, pp. 389–399.

Johnson, D. G. (1985), *Computer ethics*, 1st ed., Prentice Hall, Englewood Cliffs, New Jersey.

Johnson, D. G. (2005), 'Computer ethics', in Frey, R. G. and Heath Wellman, C. (eds.), *A companion to applied ethics*, Blackwell Publishing Ltd, Malden, Massachusetts, pp. 608–619.

Kendal, E. (2022), 'Ethical, legal and social implications of emerging technology (ELSIET) symposium', *Journal of Bioethical Inquiry*, Vol. 19, pp. 363–370.

Kesä, M. (2018), *Tekoäly tulee – muuttuuko duunarin työ?* Central Organisation of Finnish Trade Unions, Helsinki.

Krzywdzinski, M., Gerst, D. and Butollo, F. (2023), 'Promoting human-centred AI in the workplace: Trade unions and their strategies for regulating the use of AI in Germany', *Transfer*, Vol. 29, No. 1, pp. 53–70.

Lazer, D., Kennedy, R., King, G. and Vespignani, A. (2014), 'The parable of Google flu: Traps in big data analysis', *Science*, Vol. 343, pp. 1203–1205.

Leikas, J., Johri, A., Latvanen, M., Wessberg, N. and Hahto, A. (2022), 'Governing ethical AI transformation: A case study of AuroraAI', *Frontiers in Artificial Intelligence*, Vol. 5.

LO (Landsorganisationen) (2016), *Vi gillar förändring! – en progressiv agenda för framtidens jobb*, Stockholm.

LRN (2021), *LRN benchmark of ethical culture*, New York.

Manders-Huits, N. and van den Hoven, J. (2009), 'The need for a value-sensitive design of communication infrastructures', in Sollie, P. and Düwell, M. (eds.), *Evaluating new technologies: Methodological problems for the ethical assessment of technology developments*, International Library of Ethics, Law and Technology book series, Springer, Dordrecht, the Netherlands, pp. 51–60.

McKinsey Global Institute (2018), *AI, automation, and the future of work: Ten things to solve for*, briefing note.

McKinsey Global Institute (2022), *Global survey: The state of AI in 2022 – and a half a decade in review*.

Mind (2022), *Germany: Ground-breaking agreement in the retail sector signed between ver.di and H&M*, web page, accessed 4 July 2023.

Ministerio de Asuntos Económicos y Transformación Digital (2019), *Plan España Digital 2025*, Government of Spain, Madrid.

Ministerio de Trabajo, Migraciones y Seguridad Social (2019), *Resolución de 7 de octubre de 2019, de la Dirección General de Trabajo por la que se registra y publica el Convenio colectivo de Orange España*, SAU, Government of Spain, Madrid.

MIT Sloan School of Management (2021), 'Exploring the ethics of automation', blog post, 9 June.

Molina, O., Butollo, F., Makó, C., Godino, A., Holtgrewe, U., Illsoe, A. et al (2023), 'It takes two to code: A comprehensive analysis of collective bargaining and artificial intelligence', *Transfer*, Vol. 29, No. 1, pp. 1–18.

Morley, J., Floridi, L., Kinsey, L. and Elhalal, A. (2020), 'From what to how: An initial review of publicly available AI ethics tools, methods and research to translate principles into practices', *Science and Engineering Ethics*, Vol. 26, pp. 2141–2168.

OECD (Organisation for Economic Co-operation and Development) (2019), *Negotiating our way up: Collective bargaining in a changing world of work*, OECD Publishing, Paris.

Orange (2021), 'Orange creates a data and AI ethics council', press release, 23 March.

Ponce Del Castillo, A. (2020), *Labour in the age of AI: Why regulation is needed to protect workers*, Foresight Brief No. 8, European Trade Union Institute, Brussels.

Powell, A. B., Ustek-Spilda, F., Lehuedé, S. and Shklovski, I. (2022), 'Addressing ethical gaps in "technology for good": Foregrounding care and capabilities', *Big Data & Society*, pp. 1–12.

PwC (2020), *Sizing the prize. What's the real value of AI for your business and how can you capitalise?* London.

Rachovitsa, A. and Johann, N. (2022), 'The human rights implications of the use of AI in the digital welfare state: Lessons learned from the Dutch SyRI case', *Human Rights Law Review*, Vol. 22, No. 2.

Reuters (2018), *Amazon scraps secret AI recruiting tool that showed bias against women*, web page, accessed 4 July 2023.

Reuters (2023), *EU tech chief sees draft voluntary AI code within weeks*, web page, accessed 4 July 2023.

SACO (Sveriges Akademikers Centralorganisation) (2017), *Vägar till ny kunskap – konkreta förslag till en modern kompetensutveckling*, web page, accessed 18 August 2023.

SERV (Social and Economic Council of Flanders) (2018), *De transitie naar een digitale samenleving: Aanbevelingen en acties*, Brussels.

Shaw, W. (2017), *Business ethics*, 9th ed., Cengage Learning, Boston.

Smids, J., Nyholm, S. and Berkers, H. (2020), 'Robots in the workplace: A threat to – or opportunity for – meaningful work?' *Philosophy & Technology*, Vol. 33, pp. 503–522.

SPD (Sozialdemokratische Partei Deutschlands), Bündnis 90/Die Grünen and FDP (Freie Demokraten) (2021), *Mehr Fortschritt Wagen: Bündnis Für Freiheit, Gerechtigkeit und Nachhaltigkeit – Koalitionsvertrag 2021–2025*, Berlin.

Stahl, B. C. (2021), *Artificial intelligence for a better future: An ecosystem perspective on the ethics of AI and emerging, digital technologies*, Springer, Cham, Switzerland.

Swedish Work Environment Authority (2015), *Digital arbetsmiljö*. Stockholm.

Swedish Work Environment Authority (undated), *Computer work*, Solna, Sweden.

Sydsvenskan (2020), 'Tillsätt en algoritmombudsman', 3 October.

Tallo (2021), 'Tallo data insights: Gen Z's return to work plan', blog post, 14 May.

TCO (Tjänstemännens Centralorganisation) (2018), 'Den svenska partsmodellen står i centrum för teknikutveckling och omställning', press release, 27 April.

TCO (2022), 'TCO:s styrelse antar AI-utmaning', press release, 25 August.

TechGDPR (2022), *Data protection & privacy digest 2 – 15 August 2022: Commercial surveillance, sensitive data 'by comparison', worker electronic monitoring*, web page, accessed 4 July 2023.

The Economist Intelligence Unit (2020), *Staying ahead of the curve – The business case for responsible AI*, London.

The Guardian (2023), 'Big tech companies cut AI ethics staff, raising safety concerns', 29 March.

UNI Europa (2021), 'UNI Europa finance signs joint declaration with Europe's insurance employers on responsible AI', press release, 16 March.

UNI Global Union (2020), *Algorithmic management – A trade union guide*, Nyon, Switzerland.

Unionen (2020), *Ny teknik, automatisering och coronakris – så förändras tjänstemännens arbetsmarknad*, Stockholm.

United Nations (2011), *Guiding principles on business and human rights*, New York and Geneva.

Vakkuri, V., Kemell, K. K., Jantunen, M. and Abrahamsson, P. (2020), '"This is just a prototype": How ethics are ignored in software startup-like environments', in Stray, V., Hoda, R., Paasivaara, M. and Kruchten, P. (eds.), *Agile processes in software engineering and extreme programming*, Lecture Notes in Business Information Processing series, Vol. 383, Springer, Cham, Switzerland, pp. 195–210.

Vallor, S., Raicu, I. and Green, B. (2020), *Technology and engineering practice: Ethical lenses to look through*, Markkula Center for Applied Ethics, Santa Clara, California.

ver.di (Vereinte Dienstleistungsgewerkschaft) (2020), *Ethical guidelines for the development and use of artificial intelligence (AI): The common good and good work by design,* discussion paper, Berlin.

VNO-NCW (Confederation of Netherlands Industry and Employers) (2022), *Supervision of AI: AI deserves sectoral supervision coordinated by a NL AI Council*, The Hague.

Voss, E. and Bertossa, D. (2022), 'Collective bargaining and digitalization: A global survey of Union use of collective bargaining to increase worker control over digitalization', *New England Journal of Public Policy*, Vol. 34, No. 1, Article 10.

Wachter, S. (2018), 'Ethical and normative challenges of identification in the internet of things', conference presentation, *Living in the internet of things: Cybersecurity of the IoT – 2018*, 28–29 March, London.

Waller, B. N. (2011), *Against moral responsibility*, MIT Press, Cambridge, Massachusetts.

WEF (World Economic Forum) (2020), 'Don't fear AI. It will lead to long-term job growth', blog post, 26 October.

WEF (2022), 'Gen Zers are making their mark in the workplace – Here are 4 things they expect', blog post, 22 September.

WEF (2023), *Future of jobs report 2023*, Geneva.

Weller, A. (2019), 'Transparency: Motivations and challenges', in Samek, W., Montavon, G., Vedaldi, A., Hansen, L. and Müller, K. R. (eds.), *Explainable AI: Interpreting, explaining and visualizing deep learning*, Lecture Notes in Computer Science series, Vol. 11700, Springer, Cham, Switzerland, pp. 23–40.

Wolff, J. (2014), 'The precautionary attitude: Asking preliminary questions', *Hastings Center Report*, Vol. 44, No. 6, S27–S28.

Wright, D., Finn, R., Gellert, R., Gutwirth, S., Schütz, P., Friedewald, M. et al (2014), 'Ethical dilemma scenarios and emerging technologies', *Technological Forecasting and Social Change*, Vol. 87, pp. 325–336.

Yin, R. K. (2018), *Case study research and applications: Design and methods*, 6th ed., Sage Publications, Los Angeles.

Yle (2020), 'Yle Outset luopuu tekoälystä verkkokeskustelujen moderoijana – "Vain ihmissilmä pystyy hahmottamaan keskustelun sävyjä"', blog post, 20 January.

Zhang, D., Nestor, M., Brynjolfsson, E., Etchemendy, J., Lyons, T., Manyika, J. et al (2022), *The AI index 2022 annual report*, AI Index Steering Committee, Stanford Institute for Human-Centered AI, Stanford University, Stanford, California.

Annex: List of stakeholders consulted

Country	Type of stakeholder	Organisation(s)
Belgium	Government body	Belgian National Labour Council (CNT-NAR)
	Government body, employers and employees	Social and Economic Council of Flanders (SERV)
	Employers	Federation of Enterprises in Belgium (VBO-FEB)
	Employers	Réseau IA
	Employees	General Labour Federation of Belgium (ABVV-FGTB)
	Employees	Confédération des syndicats chrétiens
Denmark	Government body	Data Ethics Council
	Government body	Disruption Taskforce of the Ministry of Industry, Business and Financial Affairs
	Government body	Danish Business Authority, an agency under the Disruption Taskforce of the Ministry of Industry, Business and Financial Affairs
	Employers	Dansk Industri (DI)
	Employees	Danish Trade Union Confederation (FH)
	Employees	Ingeniørforeningen (IDA)
Finland	Government body	Ministry of Economic Affairs and Employment
	Employers	Confederation of Finnish Industries (EK)
	Employers	Finnish Confederation of Professionals (STTK)
	Employees	Central Organisation of Finnish Trade Unions (SAK)
	Employees	Industrial Union
France	Government body	Assemblée nationale
	Government body	France Stratégie
	Employers	Mouvement des entreprises de France (MEDEF)
	Employers	Fédération Syntec
	Employees	Confédération Française démocratique du Travail (CFDT)
	Employees	Union générale des ingénieurs, cadres et techniciens CGT (UGICT-CGT)
	Employees	Confédération Française de l'Encadrement – Confédération générale des cadres (CFE-CGC)
Germany	Government body	Federal Ministry of Labour and Social Affairs (BMAS) – Denkfabrik Digitale Arbeitsgesellschaft
	Government body	Ministry of Work, Health and Social Affairs of the Government of North Rhine-Westphalia
	Employers	Confederation of German Employers' Associations (BDA)
	Employers	Bundesarbeitgeberverband Chemie (BAVC)
	Employers	Südwestmetall
	Employees	German Trade Union Confederation (DGB)
	Employees	IG Metall Baden-Württemberg
	Employees	Vereinte Dienstleistungsgewerkschaft (ver.di)
Netherlands	Government body	Ministerie van Binnenlandse Zaken en Koninkrijksrelaties
	Government body	Provincie Noord-Brabant
	Employees	Christelijk Nationaal Vakverbond (CNV)
	Government body, employers and employees	Social and Economic Council (SER)
	Employees	Netherlands Trade Union Confederation (FNV)
	Employers	MKB The Hague
	Employers	Confederation of Netherlands Industry and Employers (VNO-NCW) and MKB-Nederland

Country	Type of stakeholder	Organisation(s)
Poland	Government body	Working Group on AI of the Chancellery of the Prime Minister
	Government body	Council of Ministers Committee for Digital Affairs
	Employers	Polish Craft Association
	Employers	Lewiatan Confederation
	Employees	NZSS Solidarność
	Employees	All-Poland Alliance of Trade Unions (OPZZ)
Spain	Government body	Ministry of Economic Affairs and Digital Transformation
	Government body	AI Advisory Council
	Government body	National Observatory of Technology and Society (ONTSI) of the Ministry of Economic Affairs and Digital Transformation
	Employers	Spanish Association of Digital Economy
	Employers	Multisectoral Association of Information Technology, Communications and Electronics Companies (Ametic)
	Employees	General Union of Workers (UGT)
	Employees	Fundación 1º de Mayo
Sweden	Government body	Swedish Agency for Work Environment Expertise
	Government body	Committee for Technological Innovation and Ethics
	Employees	Swedish Confederation of Professional Employees (TCO)
	Employees	Unionen
	Employers	Industriarbetsgivarna
	Employers	Confederation of Swedish Enterprise

Getting in touch with the EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://european-union.europa.eu/contact-eu_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls)
- at the following standard number: +32 22999696
- by email via: https://european-union.europa.eu/contact-eu_en

Finding information about the EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu

EU publications

You can download or order free and priced EU publications at: https://op.europa.eu/publications
Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://european-union.europa.eu/contact-eu_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: https://eur-lex.europa.eu

Open data from the EU

The EU Open Data Portal (https://data.europa.eu) provides access to datasets from the EU. Data can be downloaded and reused for free, both for commercial and non-commercial purposes.

TJ-07-23-333-EN-N

Automation and digitisation technologies, including artificial intelligence, are rapidly evolving and becoming increasingly powerful and pervasive. The full range of their effects in the workplace is yet to be seen. It is, however, important not only to explore the ethical implications of digital technologies and the effects of such technologies on working conditions as they emerge, but also to anticipate any unintended effects that raise new ethical challenges. Using a variety of research methods and building on previous research on the digital workplace, this report examines the many ramifications of digital technologies in the workplace, looking at the fundamental rights and ethical principles most at stake and the areas of working conditions most likely to be affected.

The European Foundation for the Improvement of Living and Working Conditions (Eurofound) is a tripartite European Union Agency established in 1975. Its role is to provide knowledge in the area of social, employment and work-related policies according to Regulation (EU) 2019/127.

Publications Office
of the European Union

ISBN 978-92-897-2347-3
doi:10.2806/235924