sciendo

PRODUCTION ENGINEERING ARCHIVES 2025, 31(4), 550-564

PRODUCTION ENGINEERING ARCHIVES

ISSN 2353-5156 (print)
ISSN 2353-7779 (online)

Exist since 4th quarter 2013
Available online at https://pea-journal.eu

# Artificial intelligence in industry – ethical challenges

**Izabela Marszałek-Kotzur**[1*]

[1]Silesian University of Technology, Faculty of Organization and Management, Department of Applied Social Sciences, ul. Roosevelta 26-28, 41-800 Zabrze, Poland; izabela.marszalek-kotzur@polsl.pl
*Correspondence: izabela.marszalek-kotzur@polsl.pl

**Article history**
Received 10.09.2025
Accepted 17.11.2025
Available online 12.12.2025

**Keywords**
technology,
robot,
industry,
artificial intelligence,
ethics.

**Abstract**
Artificial intelligence is revolutionising industry, significantly improving work efficiency. However, the development of AI continues to generate new ethical challenges. An international legal framework has already been developed to ensure the ethical and safe development of AI. However, it seems that new threats and ethical dilemmas continue to arise as AI develops further. This study aims to explain some of the benefits and threats of using AI in industry and to encourage reflection on the need to intensify work in this area in view of the further development of AI. Several selected ways of using AI in industry are presented, such as the implementation of cobots, algorithmic management, employee monitoring, the implementation of sustainable development ideas, and the use of generative artificial intelligence. Some of the existing legal frameworks for the safety and ethics of artificial intelligence are also described. The work is theoretical in nature and is based on a critical analysis of the literature on the subject. The reflections presented reveal a lack of consideration of the dangers associated with the responsibility involved in entrusting machines with autonomy and decision-making.

DOI: 10.30657/pea.2025.31.51

## 1. Introduction

The pace of artificial intelligence development is steadily increasing and sparking many diverse discussions, including those concerning the safety of people interacting with intelligent machines, as well as those related to ethical dilemmas. The use of AI is indispensable in almost all areas of life, and today it is hard to imagine its absence in industry, business, medicine, banking, education, marketing, communication, art, and other fields (Kuzior et al., 2019; Prokopowicz, 2023). It is essential for environmental protection (Knosala, 2023) and other activities aimed at sustainable development (Kuzior, 2020). It also plays a significant role in daily life by helping humans and making their lives easier. AI has also made a strong entry into the field of production and has permanently transformed the modern face of industry. Thanks to the use of advanced algorithms, it is possible to automate and monitor production processes with unprecedented precision. Machinery Directive. Therefore, the issue of developing ethics for artificial intelligence seems to be a necessity. The research problem presented here includes reflections on the nature of the problems arising in the process of developing artificial intelligence and its use in industry, as well as the errors that may result from this. Ethical reflection on AI is all the more important today because the use of artificial intelligence can be guided by very different goals (Von Braun et al., 2021). On the one hand, for example, it is about increasing efficiency and maximizing profits, and on the other hand, it is about greater control over human resources. The goal of AI may therefore be to increase the well-being of humans and other beings inhabiting the Earth, but it may also be the desire to manipulate on a global scale. These considerations also prompt us to ask ourselves what challenges humans face in terms of the development of ethical machines and to what extent AI ethics is possible to implement in practice. AI not only records and processes data in real time but also carries out rapid analyses, which allows for quick responses to various incidents occurring in the production process. AI algorithms analyze data while considering changing conditions and forecasts, enabling the dynamic adjustment of production processes to adapt to fluctuating demand, resource availability, and market conditions. This ability to analyze and adapt has made AI an indispensable tool in modern manufacturing. In industry, AI is used, among other things, to optimize production, automate processes, make better use of data, and other activities. Machine learning techniques and big data analysis open up new opportunities for manufacturing companies to improve various operations and increase competitiveness (Sira, 2021). Advanced algorithms are capable of analyzing vast amounts of data to optimize production schedules, minimize energy consumption, and maximize throughput. This ensures that production remains flexible, responsive, and cost-effective, even in a dynamic and uncertain environment. Artificial intelligence-based image recognition and machine learning algorithms have even revolutionized quality control processes in manufacturing. By analyzing visual data from cameras and sensors, AI systems can quickly and accurately detect defects, anomalies, and irregularities. This allows manufacturers to maintain consistently high product quality and, as a result, reduce the amount of waste generated. However, the challenges associated with its development and application, especially those of an ethical nature, require urgent solutions (Kosiak, 2023). Potential risks may arise in areas related to justice, fairness, privacy protection and data security, job loss, or the deepening of social inequalities. Questions also arise about human safety when working with robots, especially in the context of machine safety and the latest regulations, such as the Machinery Regulation. Therefore, the issue of developing artificial intelligence ethics seems to be a necessity.

The research problem presented involves reflections on the nature of issues that arise in the process of developing artificial intelligence and its use in industry, as well as the errors that may result from it. Ethical reflection on AI is particularly important today because the application of artificial intelligence can be guided by very different goals (Von Braun et al., 2021). On the one hand, it may be about increasing efficiency and maximizing profits, on the other hand about greater control over human resources. The goal of AI may therefore be the improvement of the well-being of humans and other beings inhabiting the Earth, but it may also be the desire for manipulation on a global scale. Scientists point out that AI today should be treated as part of socio-technical systems, in which, from an ethical point of view, the most important aspect is the complex web of interactions between technology and the social world. Much attention in the scientific literature is devoted to ethical challenges related to the development of artificial intelligence. Three main types of challenges related to AI are being identified: problems related to machine learning, social and political issues concerning life in a digitalized world, metaphysical questions about the nature of humans, and questions about the status of machines (Stahl, 2021). It should be noted at this point that the background of discussions on AI ethics consists of profound debates about human nature, consciousness, and mind, as well as understanding, creativity, meaning, or knowledge (Coeckelbergh, 2020). Things entrusted with trust are slowly slipping out of human control (Marszałek-Kotzur, 2024a). There are legitimate concerns that if the principles governing artificial intelligence are not properly established from the beginning, this could result in humans being subordinated to the logic of machine operation and losing control over it (Chojnowski, 2022). Humans engage with AI in various interactions based on human relationships (Heyder et al., 2023). Intelligent machines and robots may at some point begin to follow their own logic, which does not necessarily align with the assumptions and interests of their creators (Łupkowski and Jański, 2020; Wykowska, 2020; Stapels and Eyssel, 2021; Prokopowicz, 2024). Moreover, people will increasingly experience the omnipresence of technology and interactions with robots (Weng and Francesconi, 2023) as a substitute for contact with other people (Marszałek-Kotzur, 2024b). Therefore, systematic work on the continuous improvement of the existing legal standards for artificial intelligence is necessary (EC, 2021; EC, 2023; Martini 2019; Nowik 2022).

The aim of the article is to highlight a research gap resulting from the lack of sufficient protection of humans against errors made by artificial intelligence and the inability to predict certain situations that threaten health, safety, and the rights entitled to every person. The gap concerns the lack of comprehensive predictability regarding which harmful consequences for humans may ultimately arise from insufficient safeguarding of intelligent machines against unethical actions. The article also aims to demonstrate the enormous benefits of using artificial intelligence in industry while simultaneously indicating the directions from which threats may emerge. In order to identify the benefits and some errors, and consequently the risks arising from the use of artificial intelligence in industry, a thematic analysis of the literature addressing these issues was conducted.

Literature was searched in the Scopus, Web of Science, and Google Scholar databases. The thematic structure allowed for focusing on the division of materials according to the main topics, which enabled a deeper understanding of individual aspects of the study. Filters such as publication dates and document types were used, which helped narrow down the results to the most current and relevant materials, utilizing publicly available data from leading companies and scientific literature in this field. A review was conducted of research on responsible and reliable artificial intelligence in industry: Sensors 2025 – Reliability of Artificial Intelligence in Manufacturing and Computers in Industry 2024 – AI-DSS for Industry 4.0.

To identify the appropriate literature, the following search string was used: "artificial intelligence", "cobot", "predictive maintenance", "generative artificial intelligence", "industry", "ethics", "human-robot interaction", "algorithmic management", "green AI", "sustainable development", "algorithmic bias", "deskilling", "AI legal regulation" ("PN-EN ISO 10218" ISO/IEC 42001 "ISO/TS 15066", "AI Act").

The following inclusion criteria were used: peer-reviewed publications; official regulatory documents concerning AI; industrial applications of artificial intelligence in collaboration with humans (cobots), including safety and autonomy implications, ethical, social, and legal aspects; English, Polish, German languages; full-text availability; industry reports and estimates (McKinsey, Deloitte); publication years 2017-2025.

The applied exclusion criteria were: non-industrial fields; technical documents on artificial intelligence not containing discussions on ethics, safety, and legal regulations; non-reviewed blogs; popular press; unavailable full texts; items in languages other than English, Polish, or German. Texts were searched during the period from June 1, 2025, to August 31, 2025. A total of 1984 records were identified, of which 512 duplicates were removed and 1472 were examined, after which 1286 were excluded based on abstract review, leaving 186 full-text items. Ultimately, 110 items were included in the final review.

The study represents an innovative integration of research gap analyses aimed at explaining key areas of industry where the use of artificial intelligence appears necessary and essential, while simultaneously attempting to reveal potential ethical shortcomings of artificial intelligence in modern industrial plants. The research showed that existing regulations regarding human health and safety, particularly ethics, are not adequate to address the potential risks that may arise from errors caused by incorrect, unintentional, or deliberate programming and faulty autonomous data processing by intelligent, self-learning, and autonomous machines.

The methodological contribution consists in explaining some of the essential applications of artificial intelligence in modern industry and developing a framework that connects these AI applications, taking into account emerging ethical issues, with existing legal standards, enabling a systematic assessment of the accountability of implementing artificial intelligence in various organizational areas.

Due to the magnitude of the problem, the article only presents a small scope of artificial intelligence usage and its associated risks. Not all issues have been identified; only some ethically weak points of artificial intelligence application in industry have been established. Despite its shortcomings, the study may contribute to further research on the effectiveness of human rights and safety regulations in the process of applying artificial intelligence, providing well-founded conceptual frameworks that bridge the gap between reported implementation indicators and actual operational integration with humans. It could serve as an inspiration for readers engaged in ethics and the development of legal standards, as well as for those involved in the production of robots and machines endowed with artificial intelligence.

## 2. Literature review - results of research and discussion

### 2.1. Examples of artificial intelligence applications in industry

Research related to the development and application of artificial intelligence in modern industry is interdisciplinary in nature and is closely linked to studies on algorithms, cognitive technologies, and robotics (Saghiri, 2020). In the era of Industry 4.0, the convergence of digital technologies with traditional manufacturing processes is revolutionizing the industrial landscape. At the forefront of this transformation is artificial intelligence, a powerful tool that can drive innovation, optimize operations, and unlock a new level of efficiency in manufacturing. In business and industry, machine learning has already found such wide application that it is now one of the fastest-growing branches of AI (Kwilinski and Kuzior, 2019; Kwilinski et al., 2020). Their dynamic development and growing capabilities generate, besides obvious benefits, various problems, including ethical issues. The question of responsibility arises, both of the machines themselves and their creators, the problem of the status of artificial intelligence, safety in human-robot collaboration, and consequently the place of humans in the world of machines. Increasingly new legal problems are also emerging, and the issue of threatened human safety, and even its domination by artificial intelligence, is becoming more pronounced (Lubasz, 2020).

Automation and algorithmization in industry are used not only for automated tasks but also for purely intellectual tasks, until recently considered specifically human, such as reasoning, learning, and decision-making (Manroop, et al. 2024; Schaich Borg et al. 2024). The decision-making by machines involves their highly advanced autonomy (He et al., 2022). Robots and devices equipped with artificial intelligence, especially those belonging to the category of cognitive technologies, can take autonomous actions (Fragapane et al., 2022; Hongmei et al., 2022), which are difficult to predict and, although created for the benefit of humans, can cause physical, economic, or emotional harm. They also cause many inconveniences, for example psychological ones, resulting from the improper direction of technological development (Hawkins, 2023; Webb, 2019).

#### 2.1.1. Collaborative robots and human safety

An integral element of modern industry is collaborative robots, also known as cobots (Prabhakaran et al., 2024; Cohen et al., 2022; Faccio et al., 2022; Grau et al., 2023). They were written about over two decades ago (Colgate et al., 1999). However, they have only been used recently. AI-powered cobots collaborate with operators, performing repetitive or dangerous tasks with near-perfect precision and efficiency (Boschetti et al., 2023; Pizoń et al., 2023). This symbiotic relationship between humans and machines maximizes productivity, increases safety, and fosters workplace innovation, but it also poses challenges regarding human safety in contact with robots, as well as ethical challenges due to direct interaction with humans (Paliga, 2022). The main focus of researchers is both the construction of the cobot, the collaboration requirements, and how it is implemented (Paliga and Pollak, 2021). This is confirmed by publications that directly address the issue of cobots, but also treat them as part of the fourth and fifth industrial revolutions (Pizoń et al., 2023). However, it should be noted that most individual studies focus on specific aspects of cobot implementation (Boschetti et al., 2022; Grau et al., 2022), paying less attention to their evaluation (Cohen et al., 2022). Cobots differ from autonomous industrial robots because, for safety reasons, they are not isolated from humans. They have their own control systems, meaning they are not remotely controlled by a person. Cobots work in the same space as humans, so there is no physical barrier between the human and the robot. Cobots are equipped with an advanced set of sensors that monitor the environment in real time (Liu et al., 2024). By using the recorded data, they gain the ability to adapt to changing environmental conditions (Faccio et al., 2022). Their presence on production lines brings numerous benefits, supporting the production process and enhancing the capabilities and potential of company employees. Their main task is to make work easier for people by performing repetitive and monotonous tasks that can be burdensome for humans (Panagou et al., 2024). They can also help eliminate the risks associated with performing dangerous tasks. The popularity of collaborative robots stems from their ease of implementation, versatility, relatively simple software, and most importantly, the ability to work side by side with humans. Collaborative robots do not replace people but complement their work, enabling smooth and efficient operation of integrated production processes. Cobots can serve as tools that assist workers in lifting heavy objects or precisely assembling components, reducing the risk of injuries and improving work comfort. Robots can operate continuously for long hours, which increases productivity in the facility (Żywiołek, 2024). The implementation of cobots brings a major change to the manufacturing paradigm, which has so far focused on introducing new technical solutions, machines, and devices. It was economically motivated, aiming to minimize costly human labor. This change will have a significant impact both on how products are manufactured and on the open perspective of development. Combining the work of humans, machines, and IT systems into a single, coherent whole is crucial for modern industry. Therefore, in the future, further development of robotics and automation in intelligent industrial enterprises can be expected (Pizoń et al, 2023).

Despite numerous advantages, cobots face challenges related to improving operational parameters, such as increasing work speed while maintaining a high level of safety. Future models are expected to be even smarter, enabling better communication with other machines and making independent decisions regarding task execution (Schmidbauer et al. 2020). This development requires particular attention to establishing safety rules and ethical principles. Work systematically conducted by the European Parliament on introducing common regulations in the field of robotics relates not so much to defining the status of robots, but to attempting to regulate issues of accountability for their actions and ensuring human safety.

With regard to liability, it was emphasized that existing regulations concerning tort, contractual, or civil liability can indeed be applied; however, there is a need to consider the relationship between human rights, ethical and cultural references, and property rights. Taking into account the characteristics of the new generation of robots, such as intelligence, mobility, and the ability to be understood by humans, the biggest issue becomes the freedom of robots, that is, their autonomy (Kasap, 2025). Making independent decisions, along with their awareness of the environment, often leads to comparing them to human beings. The increasing level of robot autonomy necessitates the introduction of detailed regulations in the field of robotics (Marszałek-Kotzur, 2022; Lackner et al., 2024).

#### 2.1.2. Algorithmic work management and the ethics of employee performance monitoring systems in factories

Artificial intelligence and algorithms based on it increasingly directly influence the way numerous groups of employees work. Robotization, the development of digital communication and monitoring tools, and the processing of large datasets have made it possible to algorithmically describe work processes on a massive scale. The culmination of digitization processes in organizations is algorithmic management. Algorithmic management assigns tasks and monitors employee activities. Additionally, it makes decisions and evaluates performance based on indicators. It also integrates digital technologies, such as big data analysis, machine learning, geolocation, and wearable electronics, to automate or support managerial functions (Stefański and Żywolewska, 2024). Artificial intelligence plays a key role in optimizing supply chain operations, from procurement and inventory management to logistics and distribution (Bąba, 2020). Predictive analytics and machine learning algorithms help manufacturers forecast demand, optimize inventory levels, and streamline logistics processes. By optimizing the entire supply chain ecosystem, artificial intelligence enables manufacturers to reduce costs, improve efficiency, and better respond to customer needs. It also allows for multidimensional employee oversight that was previously impossible to achieve (Otto, 2022; Williams et al. 2018; Moreira, 2022). One of the more serious threats in this context is employee discrimination (Adams-Prassl et al., 2023).

The benefits of its use cannot be overestimated. Algorithmic management systems can promote employee engagement and satisfaction, but with a focus on support rather than control. For example, methods such as gamification can increase job satisfaction by creating a more engaging environment, tailoring tasks to employee preferences, and supporting cooperation between colleagues. Research has shown that when employees are involved in the task allocation process and their preferences are taken into account, they feel greater satisfaction with their work. Their productivity also improves (Kuzior and Kwilinski, 2022). However, the implementation of machine management can pose a number of challenges, such as health and safety issues, which require careful assessment and a risk mitigation plan. Functional safety risks may arise from cybersecurity threats. Automated safety control mechanisms may be vulnerable to data breaches or hacker attacks. As a result, erroneous decisions based on artificial intelligence can lead to various failures, which may expose workers to a hazardous working environment (European Agency for Safety and Health at Work, 2019). Ergonomic hazards may pose another type of risk. Sedentary work, especially in offices, increases the risk of musculoskeletal disorders, including back and neck pain (European Agency for Safety and Health at Work, 2025a).

Other potential risks are related to organizational and psychosocial issues. Algorithmic management can generate new stress factors, negatively affecting employees' mental health, job satisfaction, and well-being (Hawkins, 2023; Battal, 2025; European Agency for Safety and Health at Work, 2017). Risks may also include people losing their sense of control and autonomy. Algorithmic management systems often supervise, monitor, and control processes, limiting employee autonomy. For example, constant tracking of activities such as keystrokes, call duration, and break length leaves employees with little decision-making power (Piasna, 2024). In turn, penalties imposed on employees based on algorithmic calculations, such as automatic deductions from pay for minor delays or errors, further increase anxiety and pressure, negatively affecting employees' mental health (European Agency for Safety and Health at Work, 2025b).

In an effort to increase productivity, companies may implement systems that require employees to work without even short breaks, minimizing the time allocated to specific procedures or forcing employees to work at high speed. Competitive strategies can lead to excessive work intensification, and the algorithms used sometimes penalize employees for taking longer breaks. This increases time pressure and causes mental strain. Employees try to work faster and longer, skipping appropriate breaks (Moore, 2019). Intensive use of technology can also cause productivity to take center stage, neglecting interactions with colleagues, which can lead to feelings of loneliness and social isolation. Limited networking and a lack of integration prevent the formation of close relationships in the workplace, leading to frustration and dissatisfaction (European Agency for Safety and Health at Work, 2021; Bérastégui, 2024). If algorithmic management systems are not designed with the idea of fair treatment of all employees in mind, they can contribute to perpetuating harmful biases in hiring, promotion, and task allocation. This can result in certain groups being disadvantaged on the basis of race, gender, minority status, or other factors (Andriescu et al. 2024). Such situations can negatively affect employee morale and the organizational culture of the company. AI-based productivity assessments can also contribute to a sense of job insecurity. Employees may fear that AI-generated assessments of their work could replace human evaluation and negatively affect their promotion opportunities and job security (European Commission, 2024). Algorithms are widely used to assess employee performance in the workplace. Their valuable advantages include objectivity, scalability, data availability, and personalization. The objectivity of algorithms reduces the subjectivity present in traditional methods, which increases fairness in employee assessment. Scalability allows for the assessment of a large number of employees in a short period of time. In turn, the ability to analyze huge data sets allows for a more accurate determination of performance and identification of trends over time. The ability to adapt algorithms to the specific needs of an organization and to the characteristics of different positions allows for the personalization of results. Despite their numerous advantages, the use of algorithms in performance evaluation also has certain disadvantages. Algorithms may not take into account the specifics of the job, i.e., the context, which leads to incorrect evaluation of employees, who may be underestimated or overestimated. Dependence on data that may be incomplete or incorrect will certainly affect the results of the algorithm. In turn, a lack of knowledge about how algorithms work can lead to mistrust among employees, who may be concerned about the basis on which their assessments are made. This is known as the "black box" phenomenon (Brożek et al., 2024; Martini, 2019).

### 2.1.3. Predictive maintenance

A new stage in the development of the industry is predictive maintenance. It is based on the use of the latest technologies such as the Internet of Things, artificial intelligence, and automation. It is a strategy for maintaining machines that uses data, analysis, and artificial intelligence to predict failures and maintenance needs in machines and equipment. By using machine learning algorithms, manufacturers can analyze data from sensors and connected devices to predict equipment failures before they occur. This approach to maintenance minimizes unplanned downtime, reduces maintenance costs, and extends the lifespan of machines, which translates into lower financial losses and a reduced negative impact on the environment. As a result, overall productivity and efficiency can be improved. It minimizes the need for interventions, prevents unplanned downtime, and reduces costs. The aim of predictive maintenance is to keep physical assets in the best possible condition. The implementation of these technologies allows for increased efficiency, quality, and safety of production, as well as enhanced flexibility and scalability of manufacturing processes (Ainaz et al. 2025; Ouadah et al., 2022; Theissler at al., 2021; Sceife, 2024; Lee et al., 2023).

AI can support predictive maintenance in several ways. First, through monitoring and data analysis, AI can continuously monitor data from sensors and devices, such as vibrations, temperatures, pressure, speed, etc. Analyzing this data helps detect early signs of failures and instability in machine operations. Second, through modeling and forecasting, artificial intelligence can build models based on the analysis of historical and current data. With these models, AI can predict when machines might fail or require maintenance. Third, AI can send alerts and notifications about potential issues, allowing employees to respond to early warning signs and take appropriate action. Another way is optimizing maintenance scheduling. Thanks to data analysis, AI can optimize the maintenance schedule to minimize production downtime and costs. The next method is preventing breakdowns. Artificial intelligence can identify the causes of potential problems and suggest appropriate corrective measures, allowing failures to be avoided before they occur. Another method is decision support. Artificial intelligence can provide company employees with information and analyses that aid in making maintenance and machinery management decisions. Finally, through pattern recognition, AI can detect hidden patterns in data that may indicate the need for maintenance intervention, allowing the company to take preventive actions before more serious issues arise. With AI-supported predictive maintenance, a manufacturing company can increase machine availability, minimize maintenance and repair costs, and improve production efficiency. This allows for greater efficiency and competitiveness in the market. In this context, the appropriate infrastructure for data collection and analysis, as well as for the application of suitable technologies, is essential.

Predictive maintenance is primarily used in industries where system, machine, or component failures result in significant consequential damage and repair costs. It is mainly employed in aviation. Airlines use sensors in their machines to replace wear-resistant parts, such as turbines or pumps, at the right time to avoid ground downtime. Another area of its application is the automotive industry. Automotive companies operate some of the largest robot fleets in the world, producing tightly integrated supply chains according to specific order deadlines. Any production issue, such as an unplanned halt, would disrupt the entire supply chain. Another area of its use is wind turbines. Currently, vibration analyses can almost completely prevent wind turbine failures. Predictive maintenance is also widely used in ports, where crane failures would result in long ship waiting times and reduced port throughput. All industries use three main elements of predictive maintenance in supply chains. The first is the Internet of Things (IoT). Sensors and devices transmit data on the condition and performance of machines in real time via IoT, providing analytical systems with more reference points. The second element is software and cloud storage. Cloud computing allows for data mining or the analysis of large amounts of information using big data. The third element is predictive models. Machine learning is based on previously processed knowledge. In this way, predictive models allow for the identification of patterns, comparisons, and the creation of forecasts, which enable preparation for potential failures by planning operations before downtimes or interruptions occur (Kosicka et al., 2016). Using technology to carry out predictive maintenance of equipment allows for minimizing repair time, the cost of spare parts, and consumables, and can reduce costs by 25%, prevent failures by 70%, and increase productivity by 25% (Deloitte, 2024).

Predictive maintenance has many other advantages, such as lower energy consumption. Some failures, like overheating or compressed air leaks in the system, affect performance, causing excessive energy use. Another advantage is increased efficiency. Reducing the time spent on machine repairs, preventing their complete replacement, and shortening downtime allows planned operations to continue without unforeseen events. Predictive maintenance ensures greater workplace safety. Predicting likely events reduces the risk of accidents while maintaining personnel integrity. It also improves spare parts management. Knowing the demand for specific parts allows for optimal inventory management and ensures their availability. It is carried out only when necessary, instead of planning routine inspections, resulting in more efficient use of resources.

#### 2.1.4. Sustainable production - Green AI

An important area of AI application in industry is sustainable production (Kamiński, 2022). Sustainable production is an approach to manufacturing processes aimed at minimizing negative environmental impact while maintaining economic efficiency and ensuring social responsibility (Misztal, 2019). Implementing sustainable development principles in industry requires a comprehensive approach and innovation. One of the main aspects of sustainable production is efficient resource management. Industry strives to reduce the consumption of raw materials through recycling and the use of renewable and biodegradable materials. Moreover, by optimizing production processes, companies can significantly reduce waste and the emission of harmful substances into the atmosphere (Marszałek-Kotzur, 2023b). The help is supposed to come from the so-called green AI. This is a concept of developing artificial intelligence with the aim of minimizing its negative impact on the environment. It involves: computational efficiency; sustainable algorithm development, measuring energy impact; using renewable energy sources in data centers; reusability of models and data. AI can analyze sensor data, predict demand, and manage energy systems. The introduction of AI into industry can significantly contribute to achieving ESG-related goals. Artificial intelligence can help companies monitor and report their impact on the environment, society, and governance (Ingaldi and Ulewicz, 2025). Examples include monitoring carbon dioxide emissions, tracking sustainable production practices, and analyzing supply chain-related risks. AI also enables better resource management, which helps optimize energy and water use. Machine learning algorithms allow for more accurate forecasts regarding production and consumption, which can lead to more efficient use of raw materials and reduced waste (Wagner and Ingaldi, 2025; Kuzior et al., 2023).

However, there is a dark side to the use of AI. One of the main challenges associated with the introduction of AI into industry is rising energy costs. AI models, especially those based on machine learning, require enormous computing resources, leading to increased energy consumption. The training processes for advanced algorithms, such as language models, generate significant amounts of carbon dioxide. Although artificial intelligence can help reduce the Training large AI models requires significant computing power, which in turn consumes vast amounts of energy. Data centers that support the development and implementation of artificial intelligence are energy-intensive, contributing to environmental degradation. This paradox poses a challenge for the environmental aspect of ESG, as artificial intelligence may inadvertently contribute to the problems it seeks to solve. To reduce the negative impact of AI on the environment, technology companies must invest in more efficient energy solutions, such as sustainable data centers that use renewable energy. A key challenge in the context of ESG will be balancing the benefits of AI with minimizing its carbon footprint (European Funds, 2023). One of the key aspects of AI's social responsibility is its potential to support diversity and social inclusion (McKinsey, 2023). AI can revolutionize the labor market for people with disabilities by providing assistive tools such as voice assistants that enable communication for people with speech or mobility impairments; image recognition technologies that help blind people navigate the physical and digital world; task automation, facilitating the performance of complex operations for people with limited manual dexterity. These solutions not only improve the quality of life for people with disabilities, but also open up new career opportunities for them, which is in line with the goals of DEIB (Diversity, Equity, Inclusion, and Belonging). The convergence of artificial intelligence and environmental, social, and governance factors is a groundbreaking development in today's rapidly evolving business landscape. AI can be a catalyst for achieving ESG goals, from environmental sustainability to social responsibility to governance. When implementing ESG, companies are increasingly using AI technologies to monitor their activities for compliance with environmental, social, and governance standards. The identification of potential risks related to ESG aspects (Marszałek-Kotzur, 2023b), which can be supported by artificial intelligence, also plays an important role. AI algorithms can analyze supply chain data to identify suppliers at risk of environmental pollution. The introduction of artificial intelligence into sustainability strategies and ESG approaches can contribute to more effective management of natural resources, reduction of negative environmental impact, and increased transparency of organizations' activities. The integration of AI technology with the ESG approach opens up new opportunities in the field of sustainable development, while supporting climate change goals and promoting responsible business management on a global scale. However, the negative aspects of AI, such as carbon dioxide emissions, water consumption, increased energy consumption, and carbon footprint, cannot be ignored. The use of AI in the fight against climate factors is a natural step in the development of Industry 4.0, assuming that the ultimate environmental benefits will outweigh the potential costs (Siuta-Tokarska, 2021).

### 2.1.5. Generative artificial intelligence

A rapidly developing field of artificial intelligence in industry is generative artificial intelligence, abbreviated as GenAI. Generative artificial intelligence allows users to quickly create new content using various inputs. These inputs and outputs can include text, images, sounds, animations, 3D models, or other types of data (Burström et al., 2021; Miller, 2019). It can generate new content, texts, code, images, and even video or audio. This is made possible by advanced machine learning and deep learning algorithms. GenAI uses existing data to create unique results based on it. It has many applications in different fields. Most commonly, it is used in areas such as science and research, education, art, natural language processing, programming, entertainment, and media.In terms of technology, it is used to optimize and automate production processes by generating and testing different variants in a virtual environment. Polish researchers have developed a large language model called LongLLaMA, which has the potential to handle 64 times more text than ChatGPT. (IDEAS NCBR, 2023). The spread of generative artificial intelligence raises many concerns about ethics, potential abuse, and the quality of generated content. Because it is also trained on materials available on the internet, which are sometimes unverified, it can hallucinate, i.e., create erroneous or misleading content (Sidhu, 2024). The availability of GenAI may also create opportunities for abuse, e.g., for dishonest creation of works. There is also a risk that such tools could be used for online fraud or hacking attacks by creating credible-looking fake news or so-called deepfakes. It is also worth mentioning the social consequences.

Many experts believe that the development of GenAI could lead to the automation of many professions, which entails the risk of job losses in many industries. Furthermore, training models requires huge amounts of data, which also raises concerns about copyright infringement of existing works, creating legal risks for those who use generative AI (McKinsey, 2023). In May 2023, Google announced several new features based on generative artificial intelligence, including Search Generative Experience and a new LLM called PaLM 2, to power the Bard chatbot, as well as other Google products (Roth, 2023). Generative AI will have a significant impact on all industry sectors (McKinsey, 2023). It has enormous potential to automate work activities that currently consume 60 to 70 percent of employees' time. The acceleration of technical automation potential is largely due to the increased ability of generative AI to understand natural language, which is required for work activities that account for 25 percent of total working time (McKinsey, 2017).

The rapid development of generative AI is likely to significantly increase the overall impact of AI on industry, generating trillions of dollars in additional value each year. However, this technology may also bring new and significant challenges. The integration of machine learning mechanisms in industry ensures better resource allocation, more efficient process management, lower energy consumption, and higher safety levels. Properly implemented AI can revolutionize the way an operating system responds to the user and the environment, making it more autonomous and efficient. Generative artificial intelligence may enable labor productivity growth of 0.1 to 0.6 percent per year by 2040 (McKinsey, 2023). Given the pace of generative AI implementation to date, there is a tremendous need to accelerate digital transformation and reskill the workforce. Generative AI has the potential to create enormous value for the global economy. At the same time, it also has the potential to be more disruptive than previous generations of artificial intelligence. Generative AI, capable of the most human-like ability to use language, can be used to create misunderstandings, obscure the truth, incite violence, and even start wars.

### 2.1.6. Automated quality control, or the problem of "deskilling"

In the era of rapidly developing technologies and process automation, we increasingly encounter the concept of deskilling. Deskilling is the process of replacing human labor with automation and software. This phenomenon leads to the loss or degradation of existing skills and changes in the nature of a given profession. The causes of deskilling are seen in automation, work standardization, technological changes, changes in employment structure, and globalization. The consequences of deskilling include a decline in productivity, motivation, and job satisfaction, loss of creativity and innovation, unemployment, and social inequalities. As a result, tasks that once required highly specialized skills can now be performed with simpler training and less knowledge (Schürmann, 2025; Craddock, 2025). Thus, deskilling can have negative consequences both for individuals and for entire societies. It is a serious challenge faced by modern societies. It is a complex phenomenon that has significant implications for both employees and employers. When a machine or algorithm takes over a large part of our tasks, we stop practicing the skills we have acquired and gradually lose them.

This process already took place during the industrial revolution, when highly skilled craftsmen gave way to workers operating machines on production lines. Today, a similar phenomenon is returning in a new form. Artificial intelligence and automation are entering intellectual work, such as programming, data analysis, and content creation. Modern technologies are making human skills obsolete. Until recently, it seemed that automation mainly threatened physical labor, leaving people with tasks requiring intelligence and creativity. However, models such as ChatGPT can write essays, create code, and compose images, i.e., do what previously required human intellectAn AI revolution is coming, comparable to the one two centuries ago, but this time intellectual and creative jobs are at risk. The IT industry is at the forefront of this new revolution. Programmers are increasingly using intelligent support tools. Examples include GitHub Copilot, an "AI pair programmer" that suggests code snippets in real time, and models such as ChatGPT/Codex, which generate entire functions based on natural language descriptions. On the one hand, tedious template code writing or laborious documentation searches can be accelerated by machines. On the other hand, AI takes away the process of arriving at a solution, which was key to consolidating knowledge. When the algorithm suddenly runs out of suggestions, once routine tasks become insurmountable. Development teams also risk losing their so-called tribal knowledge. When programmers rely solely on ready-made snippets from AI, they do not consolidate their domain knowledge. If an unusual challenge arises outside of the AI's training data, without an expert, the company may become destabilized.

White-collar workers in various industries, including manufacturing, are also increasingly outsourcing part of their work to algorithms. The habit of placing trust in ready-made AI analysis can lead to the atrophy of independent work skills (Jakubiak and Stacewicz, 2023). Critical thinking and inquisitiveness may be weakened if we receive most answers directly, without our own reflection (WSJ, 2023). There is a real danger that if AI makes a subtle mistake or overlooks the context, humans may not notice it. There is a concern that human abilities may fade over time. Human intellect requires training like a muscle, and AI may relieve us of that training. The AI revolution is creating demand for new data specialists, algorithm trainers, prompt engineers, AI ethicists, and many other new professions that were previously unthinkable (Rosak-Szyrocka and Knop, 2024). Traditional professions may also see new specializations emerge. For example, programmers are learning how to formulate queries to models and verify their responses, which is becoming a valued skill (WSJ, 2023). AI can therefore be viewed not only as a threat, but also as an opportunity to raise qualifications to a new level, known as upskilling (Rafner et al. 2021). Companies report that automation has reduced the time spent on routine tasks, and the hours saved allow teams to develop skills that were not previously of interest. So how can AI be used wisely so that human skills are not lost? It seems that the key is to introduce the technology consciously. Individual users, companies, and educational institutions can take steps to counteract deskilling without giving up the benefits of AI. Whether we become mere operators of "thinking machines," losing our own proficiency, or whether we use this technology as an opportunity to develop new, higher skills, depends largely on ourselves. Instead of fighting inevitable progress, we should adapt to it wisely. This means using AI consciously, continuing to learn, and remaining creative despite the temptation of convenience (Grossman, 2023).

## 2.2. The problem of algorithmic biases

Despite their efficiency and objectivity, algorithms can be prone to errors, which may be unintended but also stem from the assumptions made by programmers. Such errors can affect how decisions are made regarding hiring, promotions, or salaries. It is important to remember that algorithms are written by programmers who make cognitive errors, lack knowledge, rely on stereotypes, or deliberately guide the algorithm toward a certain action, which can pose a serious threat. Even a minor initial error in the functioning of artificial intelligence can prove to be irreversible and costly for humans. This is called algorithmic bias. Algorithmic biases pose a serious threat to employees of organizations (Kelly-Lyth, 2021; Baeza-Yates, 2018; Danks and London, 2017).

The term "bias" is often used in psychology and refers to a tendency, inclination, or prejudice toward or against something or someone, based more on stereotypes and personal opinions than on facts and knowledge. Similarly, in relation to AI, bias means the partiality of machine learning or the partiality of algorithms and refers to systematic favoritism or discrimination. This bias occurs when AI models generate results that reflect and perpetuate human prejudices and social inequalities present in the training data or in the design of the algorithm itself. Thus, despite their efficiency and objectivity, algorithms can be prone to errors that lead to unintended biases. These biases can stem not only from the data on which the algorithms are trained, but also from the assumptions of the developers, which influence how decisions about hiring, promotions, or compensation are made (Tamir, 2025).

Algorithmic biases can have a variety of sources. For example, if employee data is biased and subjective, algorithms based on it will continue this trend. An example of a problem with bias in algorithms is a situation where a recruitment system automatically rejects female candidates for technology-related positions. This is the result of using data based mainly on male candidates from the past. The use of complex mathematical models that do not take into account all aspects of the job can lead to distorted results. The personal biases of the people creating the algorithms can also be consciously or unconsciously woven into the programming process. Biases can reinforce existing patterns of discrimination by favoring certain groups. They can cause low morale in a team when employees feel that decisions are being made unfairly. They can generate financial losses related to employee turnover or employee rights. To minimize the impact of biases, organizations should take steps to eliminate them. Eliminating bias involves implementing measures such as collecting diverse training data and regularly reviewing and analyzing the data on which algorithms are based. Training on ethics and bias is necessary, as is regular testing of algorithms for undesirable behavior, the use of diverse teams of developers representing different demographic groups to create algorithms, and the participation of experts from various fields.

The ethics of employee monitoring software refers to the ethical principles and guidelines governing the use of digital technologies to track and evaluate employee activity in the workplace. It includes the responsible and respectful use of monitoring tools to balance organizational needs with employees' rights to privacy and autonomy (Kelly-Lyth, 2021). Organizations should implement algorithms while ensuring appropriate control and transparency mechanisms. Algorithmic assessments should also be systematically supplemented with elements of human oversight and work context. Therefore, for algorithms to be fair and ethical, organizations must implement comprehensive strategies that not only minimize risk but also focus on developing a culture based on trust and transparency. Only then can the use of modern technologies serve as a tool that supports, rather than hinders, the development of teams and organizations. The European Union is introducing new regulations aimed at reducing bias in the field of artificial intelligence, with an emphasis on transparency, accountability, and ethics. The focal point is the EU AI Act, which requires companies using AI technology to assess risks. These regulations also require the introduction of principles of equality and non-discrimination into algorithms. Transparency of AI systems and industry standards are key to building trust in this technology and reducing bias. To ensure the fair and objective use of artificial intelligence, it is necessary to implement appropriate regulatory strategies (Baeza-Yates, 2018).

## 2.3. Selected documents defining the current framework for artificial intelligence management in industrial organizations

### 2.3.1. Standard PN-EN ISO 10218

In the context of the safety of robotic workstations as well as the robots themselves, one of the most important legal acts is the PN-EN ISO 10218 standard. The standard addresses issues such as: the design and construction of robots, mechanical aspects of construction, electrical and electronic aspects, behavior under various working conditions, protective measures, safeguards integrated into the robot system, emergency stop systems, markings, and warnings. Robot design must comply with safety requirements already at the design stage to minimize potential hazards (Marszałek-Kotzur, 2023a). The requirements also apply to ergonomics and maintenance, so that operators can safely work with the device. Robots must be designed in such a way as to eliminate risks associated with mechanical injuries, such as sharp edges, crushing points, or the possibility of impact. The robot's electrical systems must be properly insulated and protected to avoid electrical short circuits and surges. The robot should operate safely even under conditions such as fluctuating temperatures, humidity, or the presence of dust. This requires the use of appropriate materials and technologies. The standard also emphasizes the need for protective measures aimed at reducing the risk of human contact with hazardous parts of the robot. These measures include robot position monitoring systems that shut down the robot when people are detected nearby. Robots must be equipped with safety switches that immediately stop their operation in case of danger. They should also have clear and visible markings indicating potential hazards and required protective measures. ISO 10218 focuses on safety related both to the robots themselves and the systems around them (ISO 10218-1, 2025).

The standard specifies requirements for control and safety systems designed to ensure that the robot operates in a predictable and controlled manner. This includes the requirement that key safety functions, such as emergency stop switches, be equipped with backup systems that can also stop the robot in case the main system fails. Another requirement is the constant monitoring of the robot's position, movement speed, and other parameters to immediately detect any irregularities. It also covers technologies that enable safe human-robot collaboration, such as advanced sensors that shut down the robot if it gets too close to the operator. Robots must be equipped with mechanisms that automatically initiate appropriate safety procedures when a problem is detected. Every integrated system should have emergency functions to stop the robot or the entire system in case a hazard is detected. When humans are working alongside robots, special sensors and monitoring software should be used to automatically limit or stop the robot's movements when the operator approaches the robot's work area. In the case of collaborative robots from Universal Robots, these systems and sensors are built directly into the frame, with the possibility of additionally extending them with safety scanners that monitor the collaborative area, connected directly to the robot controller (ISO 10218-1, 2025).

### 2.3.2. TS 15066 Annex for cobots

In the context of cobots, Annex TS 15066 (ISO/TS 15066) is of significant importance. It is a technical document supplementing ISO 10218 standards and concerns human-robot collaboration in industrial environments. It was developed to facilitate safe collaboration between humans and robots in areas where robots work alongside people without additional physical protective barriers. It introduces guidelines aimed at defining safety limits in human-robot interaction, as well as guidance for designing applications with cobots. One of the most important elements of this specification is defining the permissible limits of force and pressure that can be applied to the human body in the event of a collision with a robot. The specification distinguishes four main operating modes of cobots: safe speed and separation monitoring, hand-guiding of the robot, protective space monitoring, and power and force limiting. TS 15066 provides guidelines for risk assessment for applications with cobots.

So far, work has focused on ISO 10218-1 standards concerning safety requirements for industrial robots. Part one, which provides requirements and guidance for the safe design, selection of protective measures, and information for the use of industrial robots, was published in 2011. It describes the basic hazards associated with robots and sets out requirements for eliminating or appropriately reducing the risks related to these hazards. Within the newly established Technical

Committee TC 299 Robots and Robotic Devices, work has been conducted to merge standards related to industrial and service robotics. The TC 299 Committee established six working groups focusing on different approaches and domains of standardization (EU.Robotics, 2016). ISO 10218-2 contains safety requirements for industrial robots. It specifies safety requirements concerning the integration of industrial robots and robotic systems. During the work on the ISO 10218 standard, the idea arose to prepare new standards for collaborative robots. The ISO/TS 15066:2016 document provides specific, data-driven safety guidelines necessary for risk assessment and control. It defines safety requirements for the collaboration between industrial robot systems and the work environment, as well as specific requirements and guidance for the collaboration between industrial robots and humans. This standard allows a robot and an operator to work in a shared workspace. ISO 15066:2016 meets the requirements of the European standard (ISO/TS 15066, 2016).

#### 2.3.3. Requirements for an AI management system - ISO/IEC 42001:2023 standard

On December 18, 2023, the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) published the world's first standard outlining the requirements for an artificial intelligence management system, ISO/IEC 42001:2023 (ISO, 2023). This is an event of potentially significant importance for organizations that develop, implement, or use AI-based solutions. The ISO/IEC 42001 standard specifies the requirements for establishing, implementing, maintaining, and continuously improving an AI Management System (Artificial Intelligence Management System). The structure of the standard's requirements is based on the well-known High-Level Structure (HLS), which allows its implementation to be consistent with other management system standards, such as ISO/IEC 27001, ISO/IEC 27701, or ISO 9001. This is a significant advantage for organizations already operating under integrated management systems. The standard has a systemic and universal character. It can be applied by any organization, regardless of industry, size, or technological level. Its substantive scope covers key issues specific to artificial intelligence technology, particularly the identification and assessment of AI-related risks and opportunities, the evaluation of the impact of AI systems on individuals, social groups, and the environment, AI lifecycle management – from design, through operation, to decommissioning, ensuring data quality, algorithm transparency and oversight, supervision of suppliers and technology partners, and documenting and auditing AI operations (ISO/IEC 42001, 2023).

One of the unique requirements of the standard is the obligation to conduct an AI system impact assessment, that is, an evaluation of the impact of AI systems on individuals, social groups, and society. The results of this activity should be taken into account in the risk assessment related to artificial intelligence. This approach introduces an ethical and social perspective into technology management practices. The standard was developed with real challenges related to AI in mind, such as the lack of model transparency, operational variability, risk of discrimination, or the need to ensure oversight and legal compliance. Particular emphasis was placed on the transparency, explainability, and social responsibility of AI systems (Chen, 2023).

#### 2.3.4. Framework for AI Risk Management - Act AI and Other Documents

An important document is NIST AI RMF 1.0, the AI Risk Management Framework, developed by the U.S. National Institute of Standards and Technology (NIST). It provides a comprehensive approach to identifying, assessing, mitigating, and monitoring risks associated with the design, development, deployment, and maintenance of systems utilizing artificial intelligence. It was published in January 2023 and builds on NIST's long-standing tradition in information technology risk modeling, complementing well-known standards such as the NIST Cybersecurity Framework and NIST 800-53. The NIST AI RMF describes four interconnected functions: Map, Measure, Manage, and Govern. The Map function helps define the business, legal, and social context in which the model operates. Measure focuses on the quantitative and qualitative assessment of risks, including metrics for accuracy, robustness, and environmental footprint. Manage proposes strategies for mitigating risks through appropriate technical and organizational procedures, while Govern establishes oversight principles, roles, and a continuous improvement loop covering the entire AI solution lifecycle. This structure aligns with the philosophy of Continuous Improvement and draws from the experience of the information security industry but has been adapted to the specifics of machine learning, including phenomena such as data drift, labeling errors, and threats from the model supply chain (NIST AI RMF 1.0, 2021). Other international regulations and standards include ISO 13849 and ISO 12100. These standards define the principles for designing control systems related to safety as well as general principles for risk assessment. ISO 13849 covers the design, implementation, and evaluation of the effectiveness of safety systems, while ISO 12100 provides guidelines for identifying, assessing, and reducing risk during the machine design process (ISO 13849, 2015; ISO 12100, 2010) .

In 2021, the European Commission presented a draft regulation aimed at governing this technological area, called the Artificial Intelligence Act. The AI Act explicitly recognizes the deployment of artificial intelligence systems for tasks related to algorithmic management as a high-risk context. The AI Act aims to ensure the safe and ethical use of artificial intelligence technology, taking into account citizens' rights and supporting technological innovation. The new regulations include, among other things, requirements regarding algorithm transparency, labeling of content generated by artificial intelligence, and rules for managing the lifecycle of artificial intelligence systems (Ministry of Digital Affairs, 2024). The AI Act includes a schedule of actions: from 2025, bans on unacceptable practices (Art. 5 AI Act) and the requirement to enhance AI-related skills (Art. 4 AI Act) will apply; from mid-2025 – obligations regarding general-purpose models (Arts. 52a–52d AI Act); from mid-2026, full application of provisions concerning high-risk systems will begin (Arts. 6, 9–15, 29, 61–62 AI Act) (AI ACT, 2021).

On June 13, 2024, Regulation 2024/1689 of the European Parliament and of the Council of the European Union on establishing harmonized rules for artificial intelligence was issued. The purpose of the regulation is to improve the functioning of the internal market by establishing a uniform legal framework, particularly regarding the development, placing on the market, putting into service, and use of artificial intelligence systems in the Union, in line with the values of the Union, in order to promote the dissemination of human-centered and trustworthy artificial intelligence (AI) while ensuring a high level of protection of health, safety, fundamental rights enshrined in the Charter of Fundamental Rights of the European Union, including democracy, the rule of law, and environmental protection, as well as protection from harmful effects of AI systems in the Union and supporting innovation (Official Journal of the EU, 2024).

### 2.3.5. Cybersecurity

Industrial automation is increasingly intertwined with IT networks and the Internet of Things, which brings enormous benefits but also new threats. Machines, once isolated, are now often remotely monitored, updated, and connected to the cloud – which opens the way to cyberattacks. Therefore, in European Union Regulation No. 1230 of 2023, requirements were set out aimed at machine manufacturers, integrators, and users regarding the consideration of cybersecurity principles already at the machine design stage, as well as during their operation and modernization (Regulation (EU) 2023/1230). These concern remote interference (taking control of a controller), manipulation of configurations (changing operating parameters), and triggering dangerous conditions (e.g., exceeding allowable speeds, pressures, temperatures).

Malicious digital interference can cause accidents or failures for which the machine manufacturer will be legally liable. In practice, this means the need to implement a range of new protection measures, both technical and organizational, to meet legal requirements and safeguard industrial control systems against modern threats. European Standardization Organizations, such as CEN (European Committee for Standardization) and CENELEC (European Committee for Electrotechnical Standardization), have developed harmonized standards, a set of technical standards. These standards are created on behalf of the European Commission to support European Union legislation. There is not yet a dedicated EU harmonized standard solely for machine cybersecurity, but in Machine Conformity Assessment, the manufacturer can refer to recognized standards. Machine conformity assessment is a process that not only allows machines to be legally placed on the European market, but above all ensures their safe and compliant use. The IEC/ISA 62443 family of standards, highly valued in the industry, plays a special role here. The latest part PN-EN IEC 62443-2-1:2025 describes how to establish a comprehensive security management program for control systems (Cyber Security Management System), which includes, among other things, risk analysis, policies and procedures, organizational structures, training, and continuous monitoring and improvement of security measures. Already at the machine design stage, the manufacturer is required to analyze potential cyberattack scenarios (e.g., remote takeover of the controller, sabotage of settings, ransomware blocking the HMI) and assess their impact on safety and operational continuity. Software and configuration of control systems – such as PLC controllers, SCADA systems, HMI panels, or industrial network devices – must be protected against unauthorized changes. Continuous monitoring of activity on the industrial network and devices allows for early detection of incidents. The human factor also plays a crucial role in industrial security, both as the weakest link and potentially the strongest protection, provided there is awareness of the threats. Therefore, regular training and raising awareness among all personnel is absolutely essential. The IEC 62443-2-1 standard emphasizes the importance of a clear organizational structure and training, specifying roles, responsibilities, and the required level of personnel awareness regarding cybersecurity. The goal is for every employee to understand the basics of cyber threats in their work environment and know the protective procedures (IEC/ISA 62443, 2007). Cybersecurity in industrial automation is a continuous risk management process. In addition to the above issues, questions remain regarding privacy, data ownership and security, and whether monitoring is used as a pretext for digital surveillance, performance measurement, workplace discrimination, and algorithmic management (Andriescu et al. 2024)

## 3. Summary and conclusion

Artificial intelligence is one of the key elements of industry. AI enables the automation of many processes, as well as increasing the efficiency and quality of production. AI also enables increased safety and scalability of production processes. By utilising AI-based technologies, manufacturers can unlock a new level of performance, flexibility and innovation in every aspect of the production process. From collaborative robots, algorithmic management, predictive maintenance and quality control to sustainable production and personalised production and supply chain optimisation, AI is shaping the future of industry. However, ethical considerations must be taken into account in the design and use of AI. It is necessary to identify the moral and immoral behaviours of programmers and their systems of values. Discussions about the ethics of AI are ongoing and cover various issues, but those related to ethics should not only concern the technology itself, but also humans and their condition in the present world. is why, today, we increasingly talk about AI not as something abstract or independent, but as an element of a broader social order in relation to humans. The algorithmic world of the modern human shows that machines are more efficient, effective and reliable than us, and it could be said that we don't stand a chance in a confrontation with artificial intelligence. However, it is not without its negative aspects and makes various mistakes. These errors result from both the mistakes that humans have programmed into them and their own data processing. There is an increasing awareness of the potential for harm to humans, and even the possibility of losing control over them. AI operates based on data, and data can be biased. If an algorithm "learns" prejudice, it can perpetuate it on a massive scale. The relative perfection and speed of AI often require humans to adapt to their operating principles, i.e. to act without emotion, empathy, compassion or the expression of human feelings. ack of human emotions in artificial intelligence should be corrected by incorporating ethics into it. However, ethics lie within the remit of humans, not machines. Further development of AI generates many new risks and threats, and ethics must be a priority in this development. The scale of the threats can be reduced if ethical and social issues are also considered at every stage of the process of creating, developing and implementing AI technology. This process is systematically subject to legal regulation, which should significantly reduce the likelihood of AI development getting out of human control. However, research has revealed shortcomings in this area. Current ethics still fail to address the new ethical dilemmas that arise. Technological reality therefore requires the urgent development of a new ethics, ethics adapted to the development of modern technologies and increasingly complex ethical challenges. Great technological power carries great responsibility (Marszałek-Kotzur, 2019; Fobel, 2021). Examples of unethical employee monitoring, such as intrusive surveillance or misuse of data, can lead to a lack of trust, hostility and legal consequences. e of the biggest challenges in the ethics of artificial intelligence is the broad aspect of privacy and data security. As AI increasingly enters the realm of personal and confidential information, the risk of data breaches grows. Another issue concerns the impact of AI systems on various spheres of social and political life. For example, there is a growing sense of powerlessness among employees in enterprises, who are increasingly being treated as an extension of machines. Another issue is the disappearance of various professions due to the introduction of advanced automation in different areas. The increasing automation of processes and the development of robotisation has led to a change in the demand for certain employee competencies. Most of the tasks performed by employees are routine, so their work can easily be replaced by computers and algorithms, which in turn leads to the loss of acquired human skills. Another problem is discrimination. AI relies on data that can be biased and stereotyped, leading to hurtful inferences and analyses. Discrimination is the result of algorithmic bias, which stems from training AI systems on insufficiently representative data sets. a result, artificial intelligence can take actions that are harmful and unfair towards certain groups of people (e.g. minorities). Another category of dangers associated with the development of artificial intelligence, albeit not related to errors made by AI, is the impact of technology on the environment, which cannot be overlooked here. One example is the carbon footprint resulting from the energy-intensive infrastructure required for AI systems to operate. AI ethics encompasses several issues, such as transparency and clarity of action, the use of solutions and algorithms that are free from prejudice, responsibility for potential decisions and consequences of actions, privacy and security of data. AI is a complex field that encompasses, among other things, the codification of principles, the translation of values into practical action, and critical reflection on the various economic, environmental, and socio-political contexts of AI. However, the relatively short history of AI and its dynamic development mean that we cannot treat current achievements in this field as final. People must remain true to their human values. Therefore, it is crucial to integrate ethical considerations into the process of technological innovation. AI ethics should not be viewed as an obstacle to progress, but rather as an integral element that supports sustainable development and innovation. The questions posed above, through the analysis of the literature gathered, have led us to a point where we must remind ourselves of one important philosophical question: what does it mean to be human in a world of artificial intelligence and robots? Effective use of technology is no substitute for deep reflection on reality. This article does not exhaust the topic discussed, but despite its limitations, it can serve as a starting point for further research into the development of new ethical solutions in this area. It can also encourage deeper reflection on the dual role of AI in productivity gains and potential harms to workers, opening up a discussion on ethics guidelines for production engineers.

## Reference


Adams-Prassl, J., Binns, R., and Kelly-Lyth, A., 2023. Directly discriminatory algorithms. Modern Law Review, 86, 144–175

AI Act, 2021. Proposal for a Regulation Laying Down Harmonised Rules on Articial Intelligence (AI Act) and Amending Certain Union Legislative Acts

Ainaz, J., Kim, D., Arif, M., 2025. A Survey of Predictive Maintenance Methods: An Analysis of Prognostics via Classification and Regression. DO - 10.48550/arXiv.2506.20090

Andriescu, M., Battaglini, M., Spyridopoulos, K., Kilic, L., Olausson, N., Broughton,A., Toro, D., 2024. Inteligentne cyfrowe systemy monitorowania bezpieczeństwa i higieny pracy: zastosowania i wyzwania Streszczenie. Europejska Agencja Bezpieczeństwa i Zdrowia w Pracy

Battal, S., 2025. Impact of artificial intelligence usage on occupational stress: an exploratory study of risk factors and prevention strategies. AI & Soc, DOI: 10.1007/s00146-025-02385-9

Bąba, M., 2020. Algorytmy– nowy wymiar nadzoru i kontroli nad świadczącym pracę, Praca i Zabezpieczenie Społeczne, 3

Baeza-Yates, R., 2018. Bias on the web. Communications of the ACM, 61(6), 54–61. https:// doi. org/ 10. 1145/ 32095 81

Bérastégui, P., 2024. Artificial intelligence in Industry 4.0: implications for occupational safety and health Report, No. 2024.01, ISBN 978-2-87452-714-2, European Trade Union Institute (ETUI), Brussels

Boschetti, G., Faccio, M., Granata, I., 2023. Human-Centered Design for Productivity and Safety in Collaborative Robots Cells: A New Methodological Approach. Electronics, 12(1), 167. DOI: 10.3390/ ELECTRONICS12010167

Brożek, B., Furman, M., Jakubiec, M., 2024. The black box problem revisited. Real and imaginary challenges for automated legal decision making. Artif Intell Law 32, 427–440, DOI: 10.1007/s10506-023-09356-9)

Burström, T., Parida, V., Lahti, T., and Wincent, J., 2021. AI-enabled business-model innovation and transformation in industrial eco-systems: A framework, model and outline for further research. Journal of Business Research, 127, 85–95. https:// doi. org/ 10.

Chen, Z.S. 2023. Ethics and Discrimination in Artificial Intelligence-Enabled Recruitment Practices. Humanities & Social Sciences Communications 10 (2023), 567. DOI: 10.1057/s41599-023-02079-x

Chojnowski, M., 2022. Etyka sztucznej inteligencji. Wprowadzenie, Warszawa

Coeckelbergh, M., 2020. AI Ethics, Cambridge and London.

Cohen, Y., Shoval, S., Faccio, M., Minto, R., 2022. Deploying Cobots in Collaborative Systems: Major Considerations and Productivity Analysis. International Journal of Production Research, 60(6), 1815–1831. DOI: 10.1080/00207543.2020.1870758

Colgate, E., J., Wannasuphoprasit, W., Peshkin, M., 1999. Cobots: Robots For Collaboration With Human Operators, Department of Mechanical Engineering Northwestern University Evanston
Craddock, T., 2025. AI Is Making Us Worse Programmers (Here's How to Fight Back). Medium
Danks, D., and London, A. J., 2017. Algorithmic bias in autonomous systems. In F. Bacchus & C. Sierra (Eds.), Proceedings of the twenty-sixth international joint conference on artificial intelligence, 4691–4697. International Joint Conferences on Artificial Intelligence Organization
Deloitte, 2024. State of Generative AI in the Enterprise 2024, https://www2.deloitte.com/us/en/pages/consulting/articles/state-ofgenerative-ai-in-enterprise.html, (accessed on September 19, 2025)
European Agency for Safety and Health at Work, 2017. Monitoring technology: The 21st century's pursuit of well-being?, https://osha.europa.eu/sites/default/files/Workers_monitoring_and_well-being.pdf (accessed on September 25, 2025)
European Agency for Safety and Health at Work, 2019, OSH and the future of work: benefits and risks of artificial intelligence tools in workplaces, https://osha.europa.eu/sites/default/files/202111/OSH_future_of_work artificial_intelligence.pdf, (accessed on September 25, 2025)
European Agency for Safety and Health at Work, 2021. Impact of artificial intelligence on occupational safety and health, https://osha.europa.eu/sites/default/files/Policy-brief-Impact-AI-OSH_EN.pdf, (accessed on September 25, 2025)
European Agency for Safety and Health at Work, 2025a. Musculoskeletal health and risk factors in the health and social care sector – a review of existing information Report, https://osha.europa.eu/sites/default/files/documents/musculoskeletal-health-risk-factors-hescare-sector_886_EN.pdf, (accessed on September 25, 2025)
European Agency for Safety and Health at Work, 2025b. Strategies and legislation on psychosocial risks in six european countries, https://osha.europa.eu/sites/default/files/documents/strategies-legislation-psychosocial-risks-european-countries_policy-brief_EN.pdf, (accessed on September 25, 2025)
European Commission, 2021. Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artifcial Intelligence (Artifcial Intelligence Act)
European Commission, 2023. European Declaration on Digital Rights and Principles for the Digital Decade 2023/C 23/01
European Funds, 2023. Zrównoważony rozwój firmy produkcyjnej i ochrona środowiska – ESG producenta a sztuczna inteligencja (AI 4FACTORY). Strategia ESG producenta a sztuczna inteligencja (AI) - DSR, (accessed on September 25, 2025)
Faccio, M., Granata, I., Menini, A., Milanese, M., Rossato, C., Bottin, M., Minto, R., Pluchino, P., Gamberini, L., Boschetti, G., Rosati, G., 2022. Human Factors in Cobot Era: A Review of Modern Production Systems Features. Journal of Intelligent Manufacturing, 1–22. DOI: 10.1007/S10845-022-01953-W
Fobel, P. (2021). Application of Ethical Standards a nd Ethical Risks Industry 4.0. Etyka Biznesu i Zrównoważony Rozwój. Business Ethics and Sustainable Development, 1, 20-29
Fragapane, G., Ivanov, D., Peron, M., Sgarbossa, F., Strandhagen, J.O., 2022. Increasing flexibility and productivity in Industry 4.0 production networks with autonomous. mobile robots and smart intralogistics. Annals of Operations, 308, 125-143. DOI: 10.1007/s10479-020-03526-7
Grau, A., Machado, J., Bolea, Y., Nabissi, G., Longhi, S., Bonci, A., 2023. ROS-Based Condition Monitoring Architecture Enabling Automatic Faults Detection in Industrial Collaborative Robots. Applied Sciences, 13(1), 143. DOI: 10.3390/APP13010143
Grossman, G., 2023. AI assistants boost productivity but paradoxically risk human deskilling. VentureBeat
Hawkins, T., 2023. Psychological Factors Impacting Adoption of Decision Support Tools. In: Mukherjee, S., Dutt, V., Srinivasan, N. (eds) Applied Cognitive Science and Technology. Springer, Singapore. DOI: 10.1007/978-981-99-3966-4_4
He, H., Gray, J., Cangelosi, A., Meng, Q., T., McGinnity, A., Mehnen, J., 2022. The Challenges and Opportunities of Human-Centered AI for Trustworthy Robots and Autonomous Systems, IEEE Transactions on Cognitive and Developmental Systems, vol.14, no.4, 1398-1412, 2022
Heyder, T., Passlack, N. and Posegga, O., 2023. Ethical Management of Human-AI Interaction: Theory Development Review. Journal Of Strategic Information Systems 32 (2023): 101772. DOI: 10.1016/j.jsis.2023.101772
IEC/ISA 62443, 2007, https://www.isa.org/standards-and-publications/isa-standards/isa-iec-62443-series-of-standards
IDEAS NCBR, 2023. Bardziej wydajny niż ChatGPT? Polscy naukowcy opracowali model językowy o nieznanych dotąd możliwościach, https://ideas-ncbr.pl/bardziej-wydajny-niz-chatgpt-polscy-naukowcy-longllama-model-jezykowy/. (accessed on September 25, 2025)
Ingaldi, M., Ulewicz, R., 2025. Sustainable development and technological advancements in industry 4.0: overcoming barriers in SME sector inte gration. Management systems in production engineering, 33(2), 144-162. DOI: 10.2478/mspe-2025-0015
ISO 10218-1., 2025, www.iso.org/standard/73933.html, (accessed on September 25, 2025)
ISO/TS 15066, 2016, www.iso.org/standard/62996.html, (accessed on September 25, 2025)
ISO/IEC 42001, 2023, www.iso.org/standard/42001, (accessed on September 25, 2025)
ISO 13849, 2015, www.iso.org/standard/13849, (accessed on September 25, 2025)
ISO 12100, 2010, www.iso.org/standard/12100, (accessed on September 25, 2025)
Jakubiak, M., and Stacewicz. P., 2023. Zaufanie Do Systemów Sztucznej Inteligencji. Wydanie I, Oficyna Wydawnicza Politechniki Warszawskiej. https://biblioteka.awl.edu.pl/sowacgi.php?KatID=0&typ=record&001=WROC%20W SO21000999
Kamiński R. 2022. Rozwój zrównoważony. Założenia koncepcyjne, działania systemowe i implementacja idei na poziomie przedsiębiorstwa, In: Wpływ otoczenia międzynarodowego na funkcjonowanie przedsiębiorstw w Polsce. A. Baszyński, R. Kamiński (eds). Polskie Towarzystwo Ekonomiczne w Poznaniu, Poznań, 49-79
Kasap, A., 2025. Philosophical grounds for regulating the lawful operation of autonomous vehicles, International Review of Law, Computers & Technology,1
Kelly-Lyth, A., 2021. Challenging Biased Hiring Algorithms, „Oxford Journal of Legal Studies", Vol. 41, Issue 4.
Knosala, B., 2023. Zarządzanie środowiskiem naturalnym przez sztuczną inteligencję. Ograniczenia i wyzwania narracji postnatury. Ethos 36(2023) nr 4(144) 127-145, DOI 10.12887/36-2023-4-144-10
Kosiak, P., 2023. Etyczne Aspekty Rozwoju i Wykorzystania Sztucznej Inteligencji, International Journal of New Economics and Social Sciences 5 (21). Special Issue. 341 – 350. DOI 10.5604/01.3001.0054.8421
Kosicka, E., Mazurkiewicz D., Gola A., 2016. Model adaptacyjnego algorytmu wspomagania decyzji w systemie utrzymania ruchu. Zarządzanie Przedsiębiorstwem, 19(4)
Kuzior, A., 2020. Polskie i niemieckie doświadczenia w projektowaniu i wdrażaniu zrównoważonego rozwoju, Problemy Ekorozwoju – Problems of Sustainable Development, 5(1), 81-89.
Kuzior, A., Kwilinski, A. and Tkachenko, V., 2019. Sustainable development of organizations based on the combinatorial model of artificial intelligence. Entrepreneurship and Sustainability, 7(2), 1353-1376. DOI: 10.9770/jesi.2019.7.2(39)
Kuzior, A., Kwilinski, A., 2022. Cognitive Technologies and Artificial Intelligence in Social Perception, Management Systems in Production Engi neering, 30(2), 109-115, DOI: 10.2478/mspe-2022-0014
Kuzior, A., Sira, M. and Brożek, P., 2023. Ethical Implications and Social Impacts of Integrating Artificial Intelligence into Sustainability Efforts. Zeszyty Naukowe. Organizacja i Zarządzanie / Politechnika Śląska 176 (2023): 333-46. DOI: 10.29119/1641-3466.2023.176.20
Kwilinski, A. and Kuzior, A., 2020. Cognitive Technologies in the Management and Formation of Directions of The Priority Development of Industrial Enterprises. Management Systems in Production Engineering, Volume 28, Issue 2, pp. 133-138.
Lackner, T., Hermann, J., Kuhn, C., Palm, D., 2024. Review of autonomous mobile robots in intralogistics: state-of-the-art, limitations and research gaps. Procedia CIRP, 130, 930-935. DOI: 10.1016/j.procir.2024.10.187.
Lee, S.E., Ju, N. and KH Lee, K.H., 2023, Service Chatbot: Co-Citation and Big Data Analysis toward a Review and Research Agenda. Technological Forecasting And Social Change, 194,122722. DOI: 10.1016/j.techfore.2023.122722
Liu, L., Guo, F., Zou, Z., Duffy, V.G., 2024. Application, development and future opportunities of collaborative robots (cobots) in manufacturing: A literature review. International Journal of Human–Computer Interaction, 40(4), 915-932. DOI: 10.1080/10447318.2022.2041907.

Lubasz, D., 2020. Konstruowanie humanocentrycznej SI z wykorzystaniem instrumentów prawnych ochrony danych osobowych. in: Sztuczna inteligencja w kontekście ochrony danych osobowych. UODO 2020

Łupkowski, P., Jański-Mały, F., 2020. The more you see me the more you like me: Influencing the negative attitudes towards interactions with robots. Journal of Automation, Mobile Robotics Intelligent Systems, 14(3), DOI: 10.14313/JAMRIS/3-2020/27.

Manroop, L., Malik, A. and Milner, M., 2024. The Ethical Implications of Big Data in Human Resource Management, Human Resource Management Review 34, no. 2 (2024): 101012. DOI: 10.1016/j.hrmr.2024.101012

Marszałek-Kotzur, I., 2019. Philosophical concept of responsibility - the key feature of the idea of sustainable development. Zeszyty Naukowe Politechniki Śląskiej. Organizacja i Zarządzanie, Article 135. DOI: 10.29119/1641-3466.2019.135.10

Marszałek-Kotzur, I., 2022. Cognitive technologies - are we in danger of humanizing machines and dehumanizing humans? Management Systems in Production Engineering, 30, Article 3. DOI: 10.2478/mspe-2022-0034

Marszałek-Kotzur, I., 2023a. The ethics of artificial intelligence. Can we trust robots? Zeszyty Naukowe Politechniki Śląskiej. Organizacja i Zarządzanie, Article 183. DOI: 10.29119/1641-3466.2023.183.21

Marszałek-Kotzur, I. 2023b. Raportowanie niefinansowe – raporty ESG. *Etyka Biznesu i Zrównoważony Rozwój. Interdyscyplinarne studia teoretyczno-empiryczne*, 72–79.

Marszałek-Kotzur, I., 2024a. The future of man in a world of artificial intelligence. Zeszyty Naukowe Politechniki Śląskiej. Organizacja i Zarządzanie, Article 210. DOI: 10.29119/1641-3466.2024.210.25

Marszałek-Kotzur, I., 2024b. Rozwój sztucznej inteligencji a prawa człowieka, w: Globalne konteksty poszanowania praw i wolności człowieka. Praca zbiorowa. 16. A. Kuzior (ed.), Monografia / Politechnika Śląska (No. 1048). Wydawnictwo Politechniki Śląskiej

Martini, C., 2019. Blackbox Algorithmus – Grundfragen einer Regulierung künstlicher Intelligenz., Springer Berlin, Heidelberg

McKinsey & Company, 2017. Harnessing automation for a future that works. McKinsey & Company, New York-USA https://www.mckinsey.com/featured-insights/digital-disruption/harnessing-automation-for-a-future-that-works (accessed on September 19, 2025)

McKinsey & Company, 2023. The economic potential of generative AI: The next productivity frontier, McKinsey & Company, New York-USA. https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier (accessed on September 19, 2025).

Miller, T., 2019. Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence, 267, 1–38. https:// doi. org/ 10. 1016/j. artint. 2018. 07. 007

Ministry of Digital Affairs, 2024. Europejski AI Act opublikowany, https://www.gov.pl/web/cyfryzacja/europejski-ai-act-opublikowany, (accessed on September 19, 2025)

Misztal, A., 2019. Zrównoważony rozwój przedsiębiorstw a stopień rozwoju społeczno-gospodarczego, Studia i Prace Kolegium Zarządzania i Finansów, nr 174, 33–45

Moore, P.V., 2019. OSH and the Future of Work: Benefits and Risks of Artificial Intelligence Tools in Workplaces, in: Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management. Human Body and Motion, 292-315

Moreira, T.C., 2022. Algorithms, Discrimination and Collective Bargaining. In: Collective Bargaining and the Gig Economy, DOI:10.5040/9781509956227.ch-009

NIST AI RMF 1.0, 2021, https://www.nist.gov/itl/ai-risk-management-framework, (accessed on September 19, 2025)

Nowik, P., 2022. New challenges for trade unions in the face of algorithmic management in the work environment. Studia z Zakresu Prawa Pracy i Polityki Społecznej, 29(2), 121–143.

Official Journal of the European Union, 2024. Regulation (EU) 2024/1689 of the European Parliament and of the Council

Otto, M., 2022. Dyskryminacja algorytmiczna w zatrudnieniu. Zarys problemu. Studia z zakresu Prawa Pracy i Polityki Społecznej, 29(2), 145-160. DOI:10.4467/25444654SPP.22.012.15686

Ouadah, A., Zemmouchi-Ghomari, L., Salhi, N., 2022. Selecting an appropriate supervised machinelearning algorithm for predictive maintenance. The International Journal of Advanced Manufacturing Technology,119

Paliga, M., 2022. Human–Cobot Interaction Fluency and Cobot Operators' Job Performance. The Mediating Role of Work Engagement: A Survey. Robotics and Autonomous Systems, 155. DOI: 10.1016/J.ROBOT.2022.104191

Paliga, M., Pollak, A., 2021. Development and validation of the fluency in human-robot interaction scale. A two-wave study on three perspectives of fluency. International Journal of Human-Computer Studies, 155, 102698. DOI: 10.1016/j.ijhcs.2021.102698

Panagou, S., Neumann, W.P., Fruggiero, F., 2024. A scoping review of human robot interaction research towards industry 5.0 human-centric work places. International Journal of Production Research, 62(3), 974–990. DOI: 10.1080/00207543.2023.2172473

Pizoń, J., Gola, A., Rudawska, A., 2023. Coboty jako kluczowy czynnik rozwoju masowej personalizacji w ramach Przemysłu 4.0 i Przemysłu 5.0, Inżynieria zarządzania. Cyfryzacja produkcji. Aktualności badawcze 5, 13-21), Publisher: Polskie Towarzystwo Ekonomiczne

Prabhakaran, S., Rishaa, P.S., Ashwath, R., Suresh, M., 2024. A review on cobots: A man's helping hand. AIP Conference Proceedings, 3035, 020027. DOI: /10.1063/5.0194713.

Prokopowicz, D., (2023). Opportunities and threats to the development of Artificial Intelligence applications and the need for normative regulation of this development, International Journal of Legal Studies, 2 (14), 95-130.

Prokopowicz, D., 2024. Zagrożenia wynikające z nieetycznego i nieuregulowanego prawnie zastosowania technologii sztucznej inteligencji. In: A. Wysocki (ed.), Sztuczna inteligencja: szanse i zagrożenia. Aspekty społeczne i etyczne. Warszawa

Rafner, J., Dellermann, D., Hjorth, A., Veraszto, D., Kampf, C., Mackay, W., Sherson, J., 2021. Deskilling, Upskilling, and Reskilling: a Case for Hybrid Intelligence, Morals & Machines 1(2):24-39. DOI:10.5771/2747-5174-2021-2-24

Rosak-Szyrocka, J., Knop, K., 2024. Navigating the Fourth Industrial Revolution: insights from a comprehensive bibliometric study on Industry 4.0. Production Engineering 10.30657/pea.2024.30.4

Roth, L., 2023. The nine biggest announcements from GoogleI/O, The Verge

Saghiri, A. M., 2020. A Survey on Challenges in Designing Cognitive Engines, 2020 6th International Conference on Web Research (ICWR), 165-171

Sceife, A.D.,2024. Improve predictive maintenance through the application of artificial intelligence: A systematic review. Results in Engineering, 21:101645

Schaich Borg, J., Conitzer, V., Sinnott-Armstrong, W., 2024. Moralna AI : czy bać się sztucznej inteligencji, tłum. B. Beniok, E. Ł. Łokas. Warszawa,

Schmidbauer, C., Komenda, T., Schlund S., 2020. Teaching Cobots in Learning Factories – User and Usability-Driven Implications. Procedia Manu facturing, 45, 398-404, DOI: 10.1016/j.promfg.2020.04.043

Schürmann, S., 2025. The Quiet Crisis: How is AI Eroding Our Technical Competence? DEV Community

Sidhu, B. K., 2024. Hallucinations in Artificial Intelligence: Origins, Detection, and Mitigation, International Journal of Science and Research, vol. 14, Issue 1, 8-15

Sira, M. (2021). International Aspects Of Cognitive Technology Development Based On «Sharing». Etyka Biznesu i Zrównoważony Rozwój. Interdyscyplinarne studia teoretyczno empiryczne/ Business Ethics and Sustainable Development. Interdisciplinary theoretical and empirical studies, 2, 16-22

Siuta-Tokarska, B., 2021. Przemysł 4.0 i sztuczna inteligencja: szansa czy zagrożenie dla realizacji koncepcji zrównoważonego i trwałego rozwoju? Nierówności Społeczne a Wzrost Gospodarczy, nr 65 (1/2021) Social Inequalities and Economic Growth, no. 65 (1/2021) DOI: 10.15584/nsawg.2021.1.1

Stahl, B.C., 2021. Artificial Intelligence for a Better Future. An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies. Springer Briefs in Research And Innovation Governance, https://link.springer.com/book/10.1007/978-3-030-69978-9

Stapels, J.G., Eyssel, F., 2021. Let's not be indifferent about robots: Neutral ratings on bipolar measures mask ambivalence in attitudes towards robots. PloS one, 16(1)

Stefański, K, Żywolewska, K., 2024. Lack of Transparency in Algorithmic Management of Workers and Trade Unions' Right to Information: European and Polish Perspectives. Białostockie Studia Prawnicze 29(2):53-63. DOI:10.15290/bsp.2024.29.02.05

Tamir, S., 2025. Ethical governance of AI-based humanoid carebots: the case for Ethics of Techno-care, Medicine, Health Care and Philosophy. DOI: 10.1007/s11019-025-10305-3

Theissler, A., Pérez-Velázquez, J., Kettelgerdes, M., and Elger, G., 2021. Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry. Reliability Engineering & System Safety, 215:107864

Von Braun, J., Archer, M.S., Reichberg, G.M. and Sánchez Sorond, M., 2021. AI, Robotics, and Humanity: Opportunities, Risks, and Implications for Ethics and Policy , in: J. von Braun et al. (eds), Robotics, AI, and Humanity, Springer, Cham 2021, p. 1-13.

Wagner, N., Ingaldi, M., 2025. Two faces of automation: cobots and autonomous mobile robots in action, Production Engineering Archives, 31(3), 276-290

Webb, A., 2019. The Big Nine. How the Tech Titans & Their Thinking Machines Could Warp Humanity, New York.

Weng, Y. H. and Francesconi, E., 2023. Ethical Design for Privacy-related Communication in Human-Robot Interaction, 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), Busan, Korea, Republic, 2427-2432, DOI: 10.1109/RO-MAN57019.2023.10309553

Williams, B.A., Brooks ,C.F., Shmargad, Y., 2018. How Algorithms Discriminate Based on Data They Lack: Challenges, Solutions, and Policy Implications, Journal of Information Policy, Vol. 8

Wykowska, A., 2020. Social Robots to Test Flexibility of Human Social Cognition. Int J of Soc Robotics, 12, 1203–1211. DOI: 10.1007/s12369-020-00674-5.

Żywiołek. J., 2024. Trust-building in ai-human partnerships within Industry 5.0. System Safety: Human - Technical Facility – Environment, 6(1), 89 98. DOI: 10.2478/czoto-2024-0011