Chapter 12
Digital trust and artificial intelligence: ethical standards and risk

Ewa Stawicka
Management Institute
Warsaw University of Life Sciences
Poland

Theo Anderson
Faculty of Business & Law
University of Northampton
United Kingdom

12.1. Introduction

Trustworthy Artificial Intelligence (AI) has three features that must characterize the system equipped with it throughout its life cycle: it should be legal, i.e. comply with all applicable laws and regulations, it should be ethical, ensuring compliance with ethical principles and values, and it should be robust both technically and socially, as AI systems can cause unintended harm even when used in good faith. Taken individually, each of these characteristics is necessary but not sufficient to achieve trustworthy AI. AI systems should be developed, implemented, and used in a manner consistent with the following ethical principles: respect for human autonomy, prevent harm, fairness, and explainability. You should be aware of the possible conflicts between these principles and take appropriate action in this regard. Developing trustworthy AI not only requires compliance with the law, which is only one of its three characteristics. Statutory regulations do not always keep pace with technological advances, may sometimes not align with ethical standards, or may simply not be the appropriate instrument to address specific problems. Therefore, for AI systems to be trustworthy, they should also be ethical, ensuring compliance with ethical standards. Even if an ethical goal has been met, individuals and society also need to be sure that AI systems will not cause any unintended harm. Such systems should perform their functions safely and reliably, ensuring that protective measures are provided to prevent any unintended negative effects. AI ethics is a sub-field of applied ethics that focuses on ethical issues related to the development, implementation, and use of AI. Its main purpose is to define how AI can lead to the improvement or damage to the well-being of individuals, both in terms of quality of life, as well as human autonomy and freedom, which are essential in

a democratic society. Artificial intelligence will transform many businesses and industries. Organizations most often use AI in low-risk areas, often only to derive insights from analysis. It turns out that the pace of the development of the use of artificial intelligence in practice is influenced by the lack of trust. Traditionally, trust was based on trust in family or friends, and in an extended form - for example in organizations or professional groups. Decisions made by organizations should be in line with the company's values as well as broader ethical and social standards. Stakeholder choices are based on human ethical standards and elements such as family, culture, religion, and communities. Creating a framework for using AI and managing risk may seem complicated, but the process is similar to creating controls, policies, and processes already in place for people. The risk of AI technology depends on how it is used. However, remaining under the care and control of humans presents its own challenges. The aim of the study is to assess stakeholder trust in artificial intelligence depending on ethical standards and the degree of risk. Information and data were taken from the literature sources on the subject and from researchers own study. The analysis was based on selected statistical measures and survey results, which help to answer the following research questions:

1. How is artificial intelligence assessed and used among stakeholders in Poland and Great Britain?
2. What trends can be observed in the Polish and British markets in terms of trust in AI?
3. What are the main opportunities and concerns of stakeholders related to the use of artificial intelligence?

Using Google Forms, the Authors prepared an electronic version of the survey. The resulting link was sent to several groups of people: students, employees, retirees, people of all ages.

The chapter is divided into six parts. The main problems and findings are outlined first. The second and third parts present synthetic contemporary views based on a critical review of the literature. The next chapter presents the methodology and results of the authors' research on trust in artificial intelligence among users in Poland and Great Britain. The chapter ends with a summary and research implications.

12.2 Digital Trust theoretical background

Trust is a complex, multi-faceted and multi-dimensional field of study. As a result, it is difficult to define the concept of trust in an unambiguous and universally acceptable manner (Lewis & Weigert, 1985). Trust as the basis of social relations is of interest to representatives of various disciplines of social sciences, including psychology, sociology, political science, economics, and management science. Trust is also taken care of by humanists, and above all, philosophers

and ethics (Polyanska et al., 2019). Although this concept is understood and defined in many different ways, based on the achievements of many scientific disciplines, trust can be generally defined as a kind of belief (and sometimes even certainty) that the party trusts in relation to the future behavior or state of an object (Krot et al., 2016). The tendency to trust is often considered a personality trait, hence trust can be considered both as a situational and an interpersonal variable (Ellonen et al., 2008). Objects of trust are not only people but also organizations, institutions, and enterprises that are created by people. Increasingly, however, the category of trust is considered in the context of technology development and thus refers not only to the social or institutional system but also to the technological one (Ebert, 2009). It is because gradually there is a transformation of at least some of the existing relationships between people towards the human-technology relationship. This transformation takes place as a result of an extremely dynamic, unprecedented in the world, pace of technology development, its diffusion and penetration into almost every sphere of human life. In this context, a new category of technology-oriented trust appears to be extremely interesting cognitively. Trust in technology manifests itself in human readiness to be influenced by technology, resulting from its usefulness, predictability of its effects, as well as the credibility of its suppliers (Ebert, 2009). The concept of trust, therefore, refers to the belief that the other side of the relationship, in this case, technology, will operate in a predictable and reliable manner, ensuring positive results (Łapińska et al., 2021). The inclination of an individual to use technology, is the result of one's personality traits, which is related previous experiences, openness to new experiences, willingness to constantly explore and learn (Kitsios et al., 2021). AI as the most advanced form of technology development becomes a specific challenge in the context of identification and measurement of trust. General confidence in technology is closely related to the issue of the ethics of new technologies. Ethical management in the field of technology means process transparency as well as transparency of the product itself (Wortham et al., 2017). Specifically, ethical considerations with respect to technology include: identifying potential harm, ensuring: safe design guidelines, developing security measures for new technology security, or privacy concerns (Bill et al., 2020). Another aspect relating to ethical governance technology and its impact on the general trust in technology is to ensure the confidentiality of data and information provided by the technology used. Therefore, creating data privacy policies and procedures that increase user trust in technology nowadays seems to be particularly important for building general trust in technology (Winfield et al., 2018). Artificial intelligence, as the most advanced form of technology development so far, is of interest to managers who see the possibility of increasing the competitive advantage in their applications. Generally, AI is a system or a

machine that imitates human intelligence when performing specific tasks, and can additionally improve (learn) interactively on the basis of collected information (Wang, 2019). Companies today must use the latest technology to grow and compete around the world (Lytras, Visvizi, 2021). Effective and efficient implementation of AI solutions in companies, in addition to significant financial, outlays, also requires employees' trust in its usefulness and functionality (Wilson et al., 2018).

In May 2019, the OECD (OECD, 2019) published the first intergovernmental standards for the development of Artificial Intelligence. The recommendations are intended to support innovation and at the same time build confidence in Artificial Intelligence by promoting responsible management and respect for human rights and democratic values. The principles of responsible management of Artificial Intelligence are based on the following guidelines:

AI development for prosperity and sustainable development. Artificial Intelligence should ensure human well-being and benefit the planet. Increase human capabilities, develop creativity, enable the integration of excluded groups, minimize economic, social, gender, and other inequalities. It is important that the development of technology takes place with respect for the natural environment and in accordance with the idea of sustainable development.

Concentration on values and people. The creators of Artificial Intelligence should respect the rule of law, human rights, and democratic values. Technology development must respect freedom, privacy, human dignity, data protection, justice, and social equality. Experts recommend that systems have security features that enable people to take control of the machine when the need arises.

Transparency. Artificial Intelligence systems should be designed in a transparent manner that allows the user to obtain information about the way the system works, the logic of decision making by the algorithm, and the factors influencing this decision. It is important that the user has the right to challenge the decision made by the machine.

Security. Artificial intelligence should be designed in such a way as to ensure the safety of the user, as well as prevent it from being used for inappropriate purposes. Therefore, it is essential to be able to systematically analyze the risks at each stage of the machine life cycle. Users should also be able to view the history of decisions made by algorithms.

Responsibility. The creators of Artificial Intelligence are responsible for the proper functioning of AI systems and compliance with the rules of ethics.

Trustworthy AI has three characteristics that must characterize an AI throughout its life cycle:

- it should be lawful, i.e. respect all applicable laws and regulations;
- should be ethical, ensuring compliance with ethical principles and values;

- should be robust both technically and socially, as AI systems can cause unintended harm even when used in good faith.

12.3 Ethical standards and artificial intelligence

In the area of big data with artificial intelligence, user data is transparent and very accessible, which means that it is devoid of privacy. The risk of breach of privacy is greater, and there are more victims. While AI threats continue to grow, so does the number of public and private organizations that publish ethics to guide the development and use of AI. In fact, many consider this approach the most effective proactive risk reduction strategy. Establishing ethical principles can help organizations protect individual rights and freedoms while improving well-being and the common good (European Commision, 2019). Organizations can translate these policies into standards and practices that can then be managed. Providing trustworthy AI requires not only the development of a set of rules but also the development and maintenance of an ethical culture, education, and hands-on learning. First, human dignity presupposes that every person has an "inherent value" that should under no circumstances be restricted, violated, or suppressed by others or by new technologies. In the context of AI, respect for human dignity means that all people are treated with the respect they deserve as moral entities, and not merely as objects to be screened, sorted, evaluated, collected, conditioned, or manipulated. Therefore, AI systems should be developed in a way that supports and ensures respect for and protection of the person's physical and mental integrity as well as personal and cultural identity, and guarantees that their basic needs are met (McCrudden, 2008). In the AI context, individual freedom requires the prevention of (direct) illegal coercion, threats to mental independence and mental health, unjustified supervision, misleading and unfair manipulation. AI systems must also include a commitment to ensure that their operations do not violate the fundamental obligations underpinning the rule of law and applicable laws and regulations, and to ensure procedural fairness and equality before the law. In the context of AI, equality means that system activities must not generate unfairly biased results (e.g. data used to train AI systems should be as inclusive as possible and represent different social groups). At the same time, AI applications can have a negative impact on the rights of citizens, and therefore these rights need to be protected. Experts emphasize ethical imperatives in the field of AI. These principles are (Floridi, 2018): respect for human autonomy; damage prevention; justice; possible explanation.

When developing, implementing, or using AI systems, a trustworthy AI assessment checklist should be adopted and adapted to the specific use case. Please note that this evaluation checklist will never be exhaustive. Ensuring the deployment of Trustworthy AI is about

continuously identifying requirements, evaluating solutions, delivering better outcomes throughout the entire life cycle of an AI system, and involving stakeholders in the activities undertaken. Trustworthy AI provides an excellent opportunity to support the process of countering the pressing challenges facing modern society, such as population aging, increasing social inequalities, and environmental pollution. This potential is also reflected at the global level, for example in the form of the UN Sustainable Development Goals. For example, trustworthy AI can be coupled with big data technology to ensure that energy needs can be identified more precisely, which can lead to more efficient energy infrastructure and enable more efficient use of energy. Given the aging of Europe's population, AI and robotics may prove to be valuable tools to improve the work of carers and provide support in the context of caring for the elderly and may enable real-time monitoring of patients' conditions, thus contributing to saving lives (Stawicka et al., 2021). AI ethics is widely seen as an example of applied ethics and focuses on the normative issues related to the design, development, implementation, and use of AI. Ethical AI is used to describe such development, implementation, and use of AI that ensure compliance with ethical standards, including fundamental rights, as special moral rights, as well as with ethical principles and related core values. It is the second of the three basic features necessary to achieve trustworthy AI (Guidelines, 2019). The general ethical principles of AI represent behavioral principles that are important in many cultural and geographic applications and suggest how AI solutions should behave when faced with moral decisions or dilemmas in a specific field of use. These include the principles of accountability, data privacy, and human agency. As AI and machine learning systems organically improve with the expansion of data access and the increase in computing power, they will become more effective and useful as the information ages, it continues to grow rapidly. It may not be long before AI technologies become watchdogs for supporting vital public interests and sustainable human development. The prospect that advances in AI will help mankind to face some of its most pressing challenges is: Exciting but legitimate concerns continue to abound. As with any new and rapidly evolving technology, the steep learning curve means that errors and calculations will be made and both unforeseen and damaging effects will inevitably occur. AI is no exception. In order to manage these effects responsibly and to orientate the development of AI systems towards the optimal public benefit, it will be necessary to put the ethics and security of AI in the first place. AI ethics is a set of values, principles and techniques that use commonly accepted standards of equity and inappropriate moral management in the development and use of artificial intelligence technologies (Chouldechova, 2017; Zhu et al., 2018).

In summary, the main source of the ethical problems caused by AI is the lack of ethical rules and norms. Humans can vent their emotions to a robot so that their emotions will be comforted and humans will become dependent on the robots. However, robots cannot fully meet the needs of human communication. In the long run, it is easy to cause depression, anxiety, and other psychological problems in people.

12.4 Description of the research sample

Snowball type assessment and sampling techniques were used to collect the data. The research was carried out in two populations of artificial intelligence users in Poland and Great Britain. The research was carried out at a price for Google Forms bidders. The study was conducted in late May and early September 2021 on the Internet. Using Google Forms, the authors prepared an electronic version of the survey. The link received was sent to users in Poland and Great Britain. Each of the respondents completed the questionnaire independently. The sample consisted of 180 correctly completed questionnaires. The sample in Poland consisted of 64 people (54% women, 46% men), and in the United Kingdom 116 people (32% women, 68% men). Most of the respondents have higher education - 58% in Poland and 86% in Great Britain, were employed and lived in a city with over 500,000. residents. The conducted review of the literature and the conducted analysis allowed for drawing conclusions and formulating recommendations, as well as indicated the limitations of the research.

12.5 Results

In the face of various challenges, such as: artificial intelligence and industry 4.0, new phenomena and moral dilemmas, new business models, circular economy, climate change, ethics specialists are looking for the best solutions. There are many single moral dilemmas, and new ones emerge with great intensity. Ethical standards are very often linked to cultural change. In practice, it turns out that although the vast majority of people's values and motives are common or very similar, the minority referring to cultural differences causes misunderstandings, for example in international business (Yang et al., 2014). As the definition of the concept of culture indicates, it means that culture is obvious to every human being and only getting to know it with others makes it possible to become aware of the differences. The trust aspect is also influenced by the cultural context. The research carried out on the example of Poland and Great Britain regarding the trust of users in artificial intelligence shows differences. That 64 people from Poland (54% women, 46% men) and 116 people from Great Britain (32% women, 68% men) participated in the study. Most of the respondents had higher

education, 58% of the respondents in Poland, and 86% of the respondents in Great Britain. The respondents were up to 35 years of age (42% from Poland and 28% from Great Britain), from 36 to 45 years (21% from Poland and 25% from the UK), while those over 46 are 37% from Poland and 47% from the UK.

The respondents were asked about their trust in products/services using artificial intelligence solutions. The vast majority of respondents in Poland replied that they do not trust artificial intelligence, which was answered by over 53% of respondents, while in the United Kingdom, about 24% of respondents said they did not trust AI, similarly respondents expressed their opinion about the lack of opinion on this issue.

(Figure 12.1)

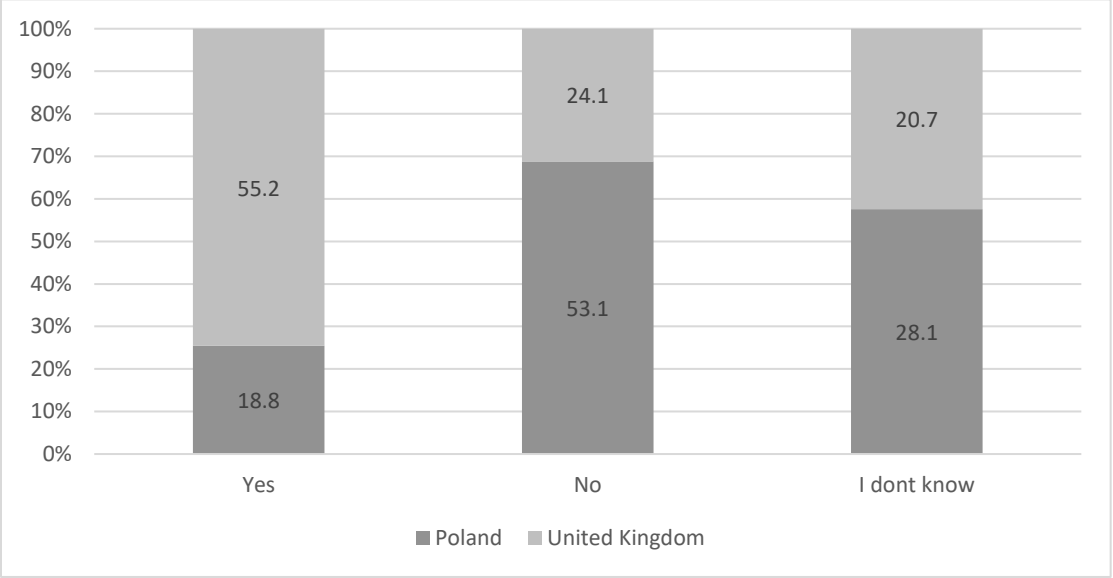

Figure 12.1. Trust to products/services that use artificial intelligence solutions by the respondents [%], N=64 Poland; N=116 United Kongdom.
Source: Author's own calculation.

As shown in Figure 12.1 trust in artificial intelligence was definitely low, especially in the case of users from Poland, although they declared that 68% of them use artificial intelligence every day, while in the case of respondents from the United Kingdom, they declared that the percentage of AI users every day was 52%. The respondents were also asked about ethical standards and risks in the use of artificial intelligence. We tried to determine what, according to users, is the greatest threat posed by products and services using artificial intelligence (the respondents could choose the three most important ones). The results are shown in Figure 12.2.

Figure 12.2

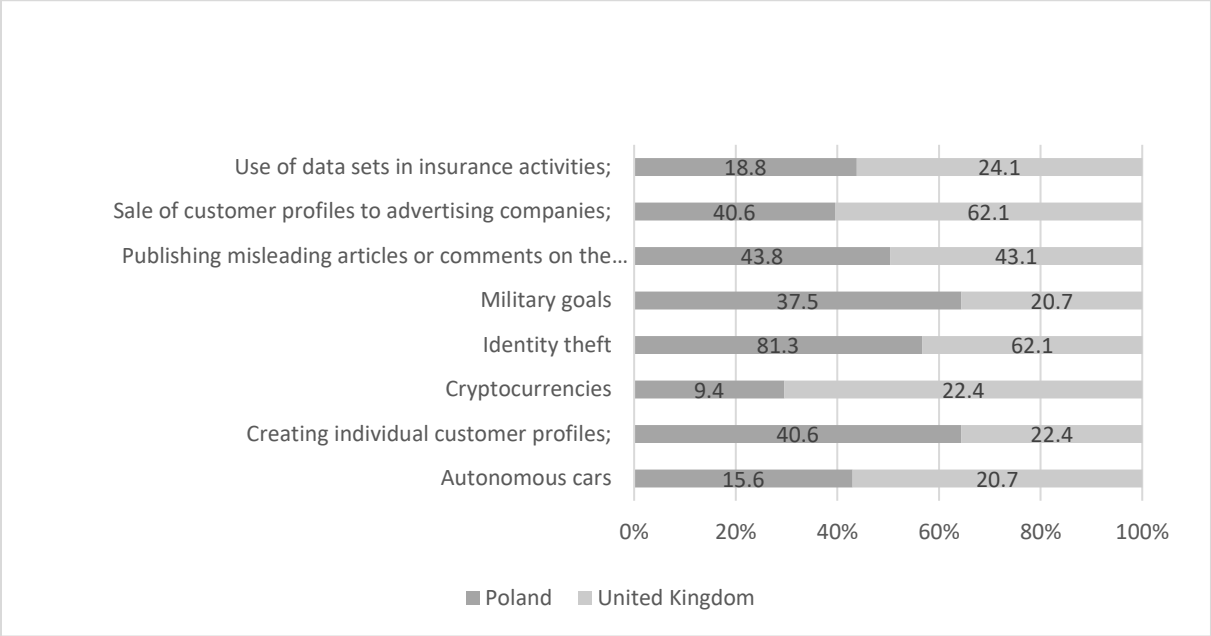

Figure 12.2. Opinion about the biggest threat posed by products and services using artificial intelligence [%], N=64 Poland; N=116 United Kongdom.
Source: Author's own calculation.

As shown in Figure 12.2, the respondents were most concerned about identity theft, as declared by over 80% of respondents in Poland and 62% of respondents in Great Britain. The next place is the fear of publishing misleading articles or comments on the Internet and violating the law or good name, which was answered by almost 44% of respondents in Poland and Great Britain. On the other hand, the third most important threat related to concerns with artificial intelligence was selling customer profiles to advertising companies, with 40% of respondents in Poland and as many as 62% of respondents in Great Britain having a negative opinion on this subject. The respondents also indicated a lack of trust and concerns about artificial intelligence related to cryptocurrency, autonomous cars as well as the use of AI for military purposes, surveillance, and stock exchange speculation. However, the last three examples were least often indicated as a threat to the respondents. The vast majority of women indicated the greatest threat in selling customer profiles to advertising companies, while among men it was indicated as dangerous, identity theft, military targets, and stock market speculation. Respondents with higher education up to 35 years of age often indicated the threat resulting from publishing misleading articles or comments on the Internet, violating the law or good name, similarly in Poland and Great Britain. The Men over 45 in Great Britain more often indicated threats in the form of Identity theft, Publishing misleading articles or comments on the Internet that violate the law or your good

name, Sale of customer profiles to advertising companies. On the other hand, women indicated Creating individual customer profiles, Identity theft, Sale of customer profiles to advertising companies. In Poland, men over 45 similarly indicated that publishing misleading articles or comments on the Internet that violate the law or good name, identity theft, and military goals. On the other hand, women in Poland over 45 indicated that they were selling customer profiles to advertising companies. In Poland, women aged 36 to 45 indicated unsafe selling of customer profiles to advertising companies, identity theft, and military targets. Women from the UK aged 36-45 indicated publishing misleading articles or comments on the Internet that violate the law or your good name, Sale of customer profiles to advertising companies, Use of data sets in insurance activities. Men from Poland pointed to selling customer profiles to advertising companies, Publishing misleading articles or comments on the Internet, violating the law or good name, Identity theft. Men from Great Britain (36-45 years old) pointed to Cryptocurrencies, Military goals, Sale of customer profiles to advertising companies. Still, both among the respondents in Poland and Great Britain, some of the respondents, when asked about trust in artificial intelligence, replied that they did not have an opinion and did not feel confident on this topic (Poland - 28.1%, Great Britain - 20.7%). This is the result of a lack of security and even less of a low level of ethics and responsibility, which largely build trust. Only strong regulations and legal provisions have a significant impact on the market, and there are still doubts about trust and ethics in the market. Lack of ethics and trust may result in even greater resistance of users to artificial intelligence.

12.6 Discussion and conclusion

This study aimed to analyze the trends that can be observed in the Polish and British-societies in terms of trust in AI. Based on the findings, several theoretical contributions were discussed. First, the need to create market confidence in artificial intelligence has been clearly identified. The attention was paid to ethical standards which should be supplemented by guidelines and safeguards. Due to the lack of moral literacy and cultural cultivation, scientists or researchers, attracted by interests, violate professional ethics and conduct illegal and ethical research, which will pose a huge threat to the interests and safety of mankind. Most of these problems can be avoided if handled properly. Nevertheless, due to the public's lack of understanding of artificial intelligence technology, it is easy to create a-misconception. For exmple, when they hear about any negative event or false information about artificial intelligence technology, they are likely

to panic, resulting in rejection (Li1 et al., 2021). First, public media should be used to create a good social atmosphere and to promote the ethics of science and technology, to help people distinguish right from wrong, to resist wrong ideas and immoral behavior, and to strengthen the public's attention to ethical values. Second, starting from the characteristics of AI-neutral technology, let the public understand that technology itself has no interests, encourage a rational approach to new things, a rational view of AI issues, so as to reduce psychological anxiety and the burden of this issue ethical. There is a need for greater stakeholder education and awareness (Yu, 2019). According to the Floridi (2016) approach, responsibility for the ethical risk of AI includes the responsibility of the designer, user, and audience. As a highly integrated product of technical attributes and social attributes, artificial intelligence public cognition is an important factor for the implementation of AI ethical norms. More ways should be expanded to enhance the public's understanding of artificial intelligence technology and inspire and cultivate the public's ethical awareness (Chen, 2020; Li et al., 2021).

The positive impact that AI systems are having on individulas and businesses will continue in economic and social terms. However, we are equally concerned with ensuring an appropriate and proportionate solution to the risks and other adverse impacts posed by these technologies in the light of the use of AI. Artificial Intelligence is a technology that is both transformative and disruptive, and its evolution over the past few years has been driven by the availability of massive amounts of digital data, major technological advances in computing power and storage capacity, and significant scientific and engineering innovations in the field of SI methods and tools. AI systems will continue to impact society and citizens in ways that are currently unimaginable. With the deep development of modern artificial intelligence, it turns out that ethics has become a problem that needs attention. Ethical deviations such as human rights, responsibility, generation, information, and prejudices caused by artificial intelligence are gradually emerging, requiring constant development and adaptation to them. The healthy and smooth development of artificial intelligence and the ethics of science and technology guidelines have a significantly positive correlation. At the same time, when preventing risk, attention should be paid to avoiding excessive limitation or the use of inappropriate methods of risk mitigation in order to inhibit the development of the industry. The article contributes to the trust literature by complementing the current debate on user trust in AI. In particular, the findings contribute to a better understanding of human collaboration and dynamics with AI, as well as the nature of human trust in AI in everyday life. Further research should be enriched with employee confidence in artificial intelligence in organizations and the impact of distrust on the performance of companies in the market.

There are some limitations that need to be identified in further research. First, the size of the sample, the test will be extended to include a larger sample. Secondly, in the future, it is worth examining the factors influencing the growth of users' trust in AI. Therefore it would be interesting to investigate how organizations could become more ethical thanks to their strategies, introduce regulations that increase trust in artificial intelligence. Ultimately, cross-sectional data were used in the study. This is necessary, and it would be interesting to use newer longitudinal data to see if the arrangements are valid over time.

References

Bill, B., Scott, B., Sandeep, S. (2020). Trends 2020. *Deloitte Insights*, 5, 1–130.

Chen, S. (2020). Algorithmic governance: risks and countermeasures of technological alienation in intelligent society. *Journal of Hubei University*. Philosophy and Social Sciences Edition, 1, 158-65.

Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big data, 5*(2), 153–163.

Ebert, T. (2009). *Trust as the Key to Loyalty in Business-to-Consumer Exchanges, Trust Building Measures in the Banking Industry*, Springer Gabler, Wiesbaden.

Ellonen, R., Blomqvist, K., & Puumalainen, K. (2008). *Journal of Innovation Management, 11*(2), 160.

Floridi, L. (2016). *The Fourth Revolution - How the Infosphere is Reshaping Human Reality*. University of Oxford.

Floridi, L. (2018). Sound Ethics and the Governance of the Digital, *Philosophy & Technology*, *31*(1), 1–8.

European Commision. (2019). *Wytyczne w zakresie etyki dotyczące godnej zaufania sztucznej inteligencji,* (Ethical guidelines for Trustworthy AI). Bruksela, 2019.

Kitsios, F., Kamariotou, M. (2021). Artificial Intelligence and Business Strategy towards Digital Transformation: A Research Agenda. *Sustainability,* 13, 2025.

Krot, K., Lewicka, D. (2016). *Zaufanie w organizacji innowacyjnej* (Trust in an innovative organization). Wydawnictwo C.H. Beck, Warszawa.

Lewis, J.D., & Weigert, A.J. (1985). *Social Forces*, *63*(4), June 1985.

Li1, H.Y., An, J.T. & Zhang, Y. (2021). *Ethical Problems and Countermeasures of Artificial Intelligence Technology.* E3S Web of Conferences 251, 01063 (2021). https://doi.org/10.1051/e3sconf/202125101063 [Accessed 17.10.2021].

Lytras, M.D., & Visvizi, A.(2021). Editorial Artificial Intelligence and Cognitive Computing: Methods, Technologies, Systems, Applications and Policy Making. *Sustainability*, 13, 3598.

Łapińska J., Escher I., Górka J., Sudolska A., & Brzustewicz P., (2021). Employees' Trust in Artificial Intelligence in Companies: The Case of Energy and Chemical Industries in Poland. *Energies*, *14*(7).

McCrudden, C. (2008). *Human Dignity and Judicial Interpretation of Human Rights*, EJIL, t. *19*(4), 655–724.

OECD (2019). *Artificial intelligence,* OECD Principles on AI. [Online] Available From: https://www.oecd.org/going-digital/ai/principles/ [Accessed 07.10.2021].

Polyanska A., Zapukhliak I., Diuk, O. (2019). *Oeconomia Copernicana, 10*(3), 561.

Stawicka E., Parlińska A. (2020). *Emerging Wireless Technologies Based on IoT in Healthcare Systems in Poland*. IoT Security Paradigms and Applications, 2020, Sudhir Kumar Sharma, Bharat Bhushan, Narayan C. Debnath, Boca Raton, CRC Press, 261-284.

Wang, P.(2019). On defining artificial intelligence. *Journal of Artificial General Intelligence,* 10, 1–37.

Winfield, A.F.T., & Jirotka, M. (2018). Ethical governance is essential to building trust in robotics and artificial intelligence systems. *Philosophical Transactions of the Royal Society,* 376, 1–13.

Wilson, H.J., Daugherty, P.R. (2018). Collaborative intelligence: Humans and AI are joining forces. *Harvard Business Review,* 96, 114–123.

Wortham, R.H., Theodorou, A. (2017). Robot transparency, trust and utility. *Connection Science*, *29*(3), 242–248.

Yu, X., Duan, W. (2019). Ethical construction of artificial intelligence. *Journal Theoretical Exploration*, 6, 9-43.

Zhu, J., Liapis, A., Risi, S., Bidarra, R., & Youngblood, G.M. (2018). *Explainable AI for designers: A humancentered perspective on mixed-initiative co-creation*. In 2018 IEEE Conference on Computational Intelligence and Games (CIG), 1-8.

Yang, J., Huiju, Y., Cen, S.-J., & Huang, R. (2014). Strategies for Smooth and Effective Cross-Cultural Online Collaborative Learning. *Educational Technology & Society*. *17*(3), 208–221.