

Research article

Ethical AI for sustainable development: User perceptions across the United Nations Sustainable Development Goals

Bo Wang^a, Jianlong Zhou^a, Fang Chen^a, Heimo Müller^b, Andreas Holzinger^{b,c}^{*}^a Human-Centered AI Lab, University of Technology Sydney, Australia^b Machine Learning and Information Science Group, Medical University Graz, Austria^c Human-Centered AI Lab, Institute of Forestry, Department of Ecosystem Management, Climate and Biodiversity, BOKU University, Vienna, Austria

ARTICLE INFO

Editor: Kuishuang Feng

Keywords:

Artificial intelligence ethics
Digital transformation
Technology in society
Machine learning
Sustainable Development Goals
Public perception

ABSTRACT

Artificial intelligence is rapidly transforming technology in society and is increasingly seen as a critical tool for addressing complex global challenges, including the United Nations Sustainable Development Goals. These seventeen goals, grouped into societal, economic, and environmental domains, present both opportunities and risks when intersecting with artificial intelligence. While artificial intelligence has the capacity to accelerate sustainable development, it may also exacerbate inequalities, environmental degradation, or other unintended harms if ethical concerns are not adequately addressed. Despite a growing body of research on ethical frameworks for artificial intelligence, there remains a lack of empirical understanding of how users perceive its potential, its ethical implications, and the principles that should guide its deployment in sustainable development contexts. It is natural to raise the questions: **How do Sustainable Development Goals and goal groups affect these user perceptions?** To answer these questions, we conducted a comprehensive human-subject study examining variations in user perceptions across 17 Sustainable Development Goals and three overarching goal groups. Our findings reveal substantial variation in perceived potential and ethical priorities depending on the specific goal, while the perceived importance of ethical considerations remains consistent across goal groups. The novelty of this study lies in combining the AI-SDG context with empirical and perception-based evidence, and our results highlight the necessity of incorporating user perspectives into the design and governance of artificial intelligence systems to ensure ethically aligned and socially accepted progress toward sustainable development.

1. Introduction

Artificial intelligence (AI) has attracted widespread interest in recent years and is increasingly regarded as a transformative force in addressing the United Nations Sustainable Development Goals (SDGs) (United Nations Department of Economic and Social Affairs, 2015). With its capacity to automate processes, generate insights, and augment human decision-making, AI holds promise for tackling complex global challenges across environmental, social, and economic domains (Vinueza et al., 2020). For example, integrating computers and automation in smart factories can advance SDG 12 (sustainable production and consumption) (Glavič, 2021). However, this optimistic narrative is counterbalanced by growing concerns over AI's unintended consequences such as ethical risks, social inequalities, and environmental externalities-arising from its behaviors becoming increasingly difficult to monitor and govern. (Zhou and Chen, 2023; Courtland, 2018). This

raises critical concerns about whether meaningful human oversight remains feasible in such systems, as explored by recent scholarship questioning the very possibility of maintaining effective human-in-the-loop control in high-autonomy contexts (Holzinger et al., 2025). This raises a critical issue that has garnered significant scholarly and public discourse: AI ethical concerns, including the growing difficulty of maintaining effective human oversight as systems become more autonomous and opaque.

AI ethical concerns, such as privacy leakage, discrimination, unemployment, and security risks, brought about by AI systems have posed significant challenges to individuals and society (Zhou et al., 2020; Huang et al., 2022; Zhou et al., 2024). Consequently, growing debates have emerged on the need for AI ethics to guide the responsible development and deployment of AI technologies, ensuring alignment with human and societal values (Lo Piano, 2020; Gupta et al., 2020).

^{*} Corresponding author at: Human-Centered AI Lab, Institute of Forestry, Department of Ecosystem Management, Climate and Biodiversity, BOKU University, Vienna, Austria.

E-mail address: andreas.holzinger@boku.ac.at (A. Holzinger).

<https://doi.org/10.1016/j.spc.2025.09.014>

Received 3 July 2025; Received in revised form 27 September 2025; Accepted 28 September 2025

Available online 8 October 2025

2352-5509/© 2025 The Authors. Published by Elsevier Ltd on behalf of Institution of Chemical Engineers. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

AI ethics—a field dedicated to studying the principles and guidelines related to AI—has been widely discussed as a means to make AI uphold ethical norms and behave in an ethical manner (Siau and Wang, 2020). To this end, organizations have proposed numerous ethical principles like transparency and fairness to mitigate and address emerging ethical concerns (The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, 2017; Floridi et al., 2018; Jobin et al., 2019; Angerschmid et al., 2022). When rigorously implemented, such principles enable AI to operate ethically to advance progress toward the SDGs.

With the advancement of ethical AI, the intersection of AI, AI ethics, and SDGs has gained increasing attention. Given that both the development of ethical AI and the pursuit of the SDGs are fundamentally human-centric and aimed at societal benefit, the involvement of the public perceptions is critical to aligning technological implementation with collective values (Ryan et al., 2020; Kieslich et al., 2022). Furthermore, understanding public perceptions of ethical principles can shed light on the design and deployment of ethical AI systems aligned with SDGs (Berendt, 2019). Therefore, it becomes imperative to assess how SDGs affect public perceptions of AI, AI ethics, and most major AI ethical principles. As the SDGs are categorized into society, economy, and environment groups (Vinueza et al., 2020; Stockholm University, 2017), it is equally critical to evaluate these perceptions across different SDG groups. Disparities in these perceptions can inform the development of contextually tailored AI systems to support the effective delivery of the SDGs. While previous research has explored the intersection of AI, AI ethics, and SDGs (Yeh et al., 2021), it has not systematically investigated public perceptions of AI across SDGs groups, nor assessed perceptions of AI ethical principles when AI is applied to different SDGs. Driven by these gaps, our investigation aims to answer the following research questions:

RQ1: How do users' perceptions of AI's potential, its ethical implications, and the prioritization of specific AI ethical principles vary across SDGs in the context of AI-driven sustainable development?

RQ2: How do users' perceptions of AI's potential, its ethical implications, and the prioritization of specific AI ethical principles differ across goal groups in the context of AI-driven sustainable development?

In response to **RQ1**, we design a human-based experimental study with 136 participants, introducing the SDGs scenario where participants rate their perceptions of the potential of AI and the role of AI ethics and make decisions of their perceived most important AI ethical principles for each of the 17 SDGs. In this study, the individual SDGs were treated as the independent variable. While our primary analysis examines the impacts of individual SDGs on users' perceptions, we also address **RQ2** through a post-hoc analysis by grouping the SDGs into three groups (society, economy, and environment). Specifically, these groups are not presented to participants during the study and are introduced during the analysis phase to investigate whether user perceptions vary across these groups. Moreover, given the need for holistic ethical frameworks in AI-driven SDG initiatives, we draw on eleven globally recognized principles identified by Jobin et al. (2019). The main contributions of this paper are as follows:

- We conduct an online user study to evaluate the effects of SDG and SDG groups (society, economy, environment) on user perceptions of AI's potential, the role of AI ethics, and the most important AI ethical principle.
- We, to the best of our knowledge, are the first to present novel empirical evidence on user perceptions of specific AI ethical principles within the context of AI-driven SDG initiatives.
- We develop and employ three subjective measures to assess user perceptions of the potential of AI, the role of AI ethics, and the most important AI ethical principle.

2. Related work and hypotheses

In this section, we first examine the literature on the potential of AI, the development of AI ethics, and its role in advancing SDGs. Subsequently, we formulate our hypotheses considering different SDGs and SDG groups.

2.1. Related work

2.1.1. AI and SDGs

AI with powerful capabilities in prediction, automation, and personalization is transforming our world and influencing nearly every aspect of modern life (Zhou and Chen, 2018). This potential has spurred initiatives in AI for Social Good (AI4SG), which aim to leverage the capabilities of AI to deliver positive social impacts (Cowls et al., 2021; Tomašev et al., 2020). Such advancements position AI as an enabler of the SDGs, addressing existing limitations. For instance, AI-driven innovations such as electric autonomous vehicles and smart appliances can benefit SDGs 7 (affordable and clean energy), 11 (sustainable cities and communities), and 13 (climate action) by optimizing energy efficiency and reducing urban carbon footprints (Fuso Nerini et al., 2019; Agency, 2017; Ryan et al., 2020). Besides, looking at SDG 12, there is evidence that AI-enabled technologies reduce the costs of production, enhance productivity, and deepen automation (Acemoglu and Restrepo, 2018). Furthermore, AI advances can serve as enablers for SDGs 14 (life below water) and 15 (life on land) (Keramitsoglou et al., 2006; Mohamadi et al., 2016). By leveraging satellite images, AI algorithms can automatically detect oil spills to mitigate marine pollution and classify vegetation types to combat desertification. While studies highlight AI's potential to advance all Sustainable Development Goals (SDGs), its deployment also has negative impacts—even within the same SDG (Vinueza et al., 2020). For example, although AI can enable companies to manage increasing supply chains more effectively, this may conflict with the shorter value chains necessary to achieve SDG 12 (Sadowski and Powell-Tuck, 2017). There is another vital shortcoming of AI in the context of SDGs 7 and 13. A notable example is Bitcoin mining, whose resource-intensive nature results in excessive electricity consumption and high emissions, undermining affordable energy and climate mitigation efforts (Truby, 2018). Moreover, AI can hinder SDG 5 (gender equality) as word embeddings have been shown to exacerbate existing gender stereotypes (Bolukbasi et al., 2016). Also, AI-driven automation systems may exacerbate socioeconomic inequalities, counteracting progress toward SDG 10 (reduced inequalities) by widening wage gaps between less-educated and more-educated workers and diminishing opportunities for labor market inclusion of low-skilled workers (Vinueza et al., 2020; Bundy, 2017; Nagano, 2018). In addition, AI applied to agricultural techniques may not be accessible to small farmers, potentially bringing negative effects on SDG 2 (zero hunger) (Vinueza et al., 2020). These academic findings confirm AI's dual role as both an enabler and an inhibitor in SDGs. In this paper, we focus on the investigation of the potential of AI in advancing SDGs from user perspectives.

2.1.2. AI ethics and SDGs

The field of AI ethics, particularly in contexts advancing the SDGs and broader societal well-being, seeks to systematically examine and address emerging ethical issues in the development and deployment of AI systems (Astobiza et al., 2021). As AI becomes more prevalent alongside its potential adverse societal impacts, the need for implementing principles of regulation and control of AI becomes increasingly critical (Vinueza et al., 2020; Astobiza et al., 2021; Zhou and Chen, 2023). In response, there are many ethical guidelines have been proposed to navigate risks and guide the responsible integration of AI. For instance, the Institute of Electrical and Electronics Engineers (IEEE) has issued an Ethically Aligned Design report to mitigate a wide range of

ethical concerns such as prioritizing the maximum benefit to humanity (The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, 2017). In addition, Floridi et al. (2018) suggest an emerging principle—*explicitability*—understood as incorporating both intelligibility and accountability, supporting for the establishment of a Good AI Society. Furthermore, Jobin et al. (2019) have identified a set of 11 widely recognized ethical principles through a systematic overview of 84 ethical guidelines worldwide: transparency, justice and fairness, non-maleficence, responsibility, privacy, beneficence, freedom and autonomy, trust, sustainability, dignity, and solidarity. These globally convergent principles represent the most recurrent and internationally recognized principles, thus offering a foundational framework for aligning AI ethics with the goals of sustainable development. In this paper, we select this set of principles due to their broad prevalence across major international AI ethics guidelines.

Recent studies have discussed the role of AI in achieving SDGs (Variale et al., 2024), what is still missing is to expand this for AI ethics. While certain principles, such as transparency, have been identified as crucial for minimizing harm and enhancing AI's contribution to achieving SDG 11 (Pastor-Escuredo et al., 2022; Society, 2017). Besides, fairness is closely tied to certain SDGs such as SDG 8 (decent work and economic growth) and SDG 10 (Vinuesa et al., 2020; Bourhis et al., 2019; Bankins and Formosa, 2023). Specifically, unfair compensation or biased job allocation can negatively impact SDG 8, while unequal access to AI technologies in low- and middle-income countries may exacerbate inequalities, hindering progress on SDG 10. Additionally, Ryan et al. (2020) highlight the relevance of privacy and non-maleficence to SDG 7. They conclude that end-users may have privacy concerns when engaging smart information systems in the electricity energy sector, and the complexity of decentralized architectures increases vulnerability to cyberattacks, leading to energy disruptions. Also, sustainability is linked to the achievement of SDG 12 (Amodei et al., 2016). They have argued that companies deploying AI to address specific environmental tasks, while neglecting broader environmental considerations, will be hindered in adopting truly sustainable practices. So far, there is little work exploring the role of AI ethics when using AI to achieve SDGs from user perspectives. In this paper, we investigate the perceived role of AI ethics and perceived most major AI ethical principle—among 11 principles concluded by Jobin et al. (2019)—in AI-driven SDG initiatives.

2.2. Hypotheses

To answer our research questions (R1 and R2), we formulate the following hypotheses (H1 and H2), considering different SDGs (17 SDGs) and various SDG groups (society, economy, and environment).

H1: When using AI in SDGs, user perceptions of AI's potential, the role of AI ethics, and the most important ethical principle will be influenced by various SDGs. Specifically, we expect that various SDGs will result in different levels of perceived AI's potential and perceived role of AI ethics, as well as different perceived most critical principles. For H1, we have three sub-hypotheses:

- H1.1: The perceived potential of AI to achieve SDGs will be distinct for various SDGs.
- H1.2: The perceived role of AI ethics in using AI to achieve SDGs will be distinct for various SDGs.
- H1.3: The perceived most important AI ethical principles will be distinct for various SDGs.

H2: In the context of AI-driven SDG initiatives, user perceptions of AI's potential, the role of AI ethics, and the most important ethical principle are expected to vary across different SDG groups. Specifically, we anticipate that various SDG groups will be associated with distinct levels of perceived AI's potential and perceived role of AI ethics, and different perceived most critical principles. For H2, we have three sub-hypotheses:

- H2.1: The perceived potential of AI to achieve SDGs will be different for various SDGs per different SDG groups of society, economy, and environment.
- H2.2: The perceived role of AI ethics in using AI to achieve SDGs will be different for various SDGs per different SDG groups of society, economy, and environment.
- H2.3: The perceived most important AI ethical principles will be different for various SDGs per different SDG groups of society, economy, and environment.

Hypothesis H1 focuses on the effects of different SDGs on user perceptions, while H2 pertains to the impacts of various SDG groups on user perceptions.

3. Methodologies

We begin with an overview of the study design, followed by a discussion of the approach used to measure user perceptions. Subsequently, we present participant demographic information and conclude with a detailed description of the overall study procedure.

3.1. Study design

To address our two RQs, we conduct an online user study employing a repeated measures design within the context of AI-driven SDG initiatives. The design is particularly powerful, as it allows for robust statistical analysis with a limited sample size by requiring each participant to respond to all items (Rietveld and Van Hout, 2007). In this study, tasks are designed to be straightforward and easy to ensure accessibility for non-expert participants, requiring them to evaluate three dimensions: the perceived potential of AI, the perceived role of AI ethics, and the perceived most important AI ethical principle. To test our hypotheses, we first have a primary analysis to examine the effects of different SDGs on user perceptions. Additionally, we conduct a post hoc analysis to explore variations in user perceptions across SDG groups.

In this study, the primary independent variable comprises the 17 SDGs, treated as distinct experimental conditions to isolate their individual effects on user perceptions. In addition, to examine broader thematic patterns, SDGs are post hoc grouped into three categories: economic, society, and environment. Specifically, while participants are unaware of these groupings during the study, these groupings are applied only during the analysis phase to examine variations in user perceptions across SDG groups. Table 1 presents the three categorized groups of SDGs along with all 17 SDGs. By considering different SDGs, 17 tasks are conducted by each participant in the user study. To mitigate potential order effects inherent in a repeated measures design, the task sequence is randomized for each participant. We conduct a pilot test with eight participants to ensure the survey length and cognitive load are manageable, leading to an optimized final study design and improved question presentation.

3.2. Measures

We have employed three subjective measures to evaluate user perceptions of AI and AI ethics across different SDGs and SDG groups. These three measures probe participants' perceived potential of AI, perceived role of AI ethics, and perceived most important AI ethical principle, respectively. The complete set of statements is listed in Table 2.

Perceived potential of AI. We evaluate the human perceived potential of AI for each SDG using one statement derived from the propositions of Hoffman et al. (2018); for instance, 'Do you think that AI can end poverty?'. Participants are asked to rate this statement on a 5-point Likert scale for each SDG from a range of 1 (highly unlikely) to 5 (highly likely), quantifying perceived AI's capacity to drive progress within each SDG context.

Table 1
Categorization of the SDGs into three groups.

Group	SDGs
Society	SDGs 1 (No poverty), 2 (Zero hunger), 3 (Good health and well-being), 4 (Quality education), 5 (Gender equality), 6 (Clean water and sanitation), 7 (Affordable and clean energy), 11 (Sustainable cities and communities), 16 (Peace, justice and strong institutions)
Economy	SDGs 8 (Decent work and economic growth), 9 (Industry, innovation and infrastructure), 10 (Reduced inequalities), 12 (Responsible consumption and production), 17 (Partnerships for the goals)
Environment	SDGs 13 (Climate action), 14 (Life below water), 15 (Life on land)

Table 2
Three subjective statements for measuring users’ perceptions of AI and AI ethics.

Measure	Statement	Type
Perceived role of AI	Do you think that AI can achieve [SDG]?	5-point Likert scale
Perceived role of AI ethics	Do you think AI ethics should be considered when using AI in [SDG]?	5-point Likert scale
Perceived most important AI ethical principle	Which is the most major AI ethical principle that should be considered when using AI in [SDG]?	Multiple-Choice Questions

Perceived role of AI ethics. Along similar lines, to gauge the human perceived role of AI ethics for each SDG when using AI in SDGs, one statement is developed from the measurements proposed by Hoffman et al. (2018); for example, ‘Do you think AI ethics should be considered when using AI in ending poverty?’. This statement utilizes a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree).

Perceived most important AI ethical principle. Inspired by the methods adopted by Ribeiro et al. (2016) and Zheng et al. (2023), we generate one statement to evaluate the perceived most important AI ethical principle for each SDG; for instance, ‘Which is the most important AI ethical principle that should be considered when using AI in ending poverty?’. Specifically, this statement asks participants to determine the most important AI ethical principle for each SDG from 11 principles identified by Jobin et al. (2019): transparency, justice & fairness, non-maleficence, responsibility, privacy, beneficence, freedom & autonomy, trust, sustainability, dignity, and solidarity.

3.3. Sample

The study had 136 participants, invited through social networks and direct mailing lists to ensure a wide reach and diverse responses. Among the participants, 73 identified as female, 53 as male, 6 preferred not to disclose their genders, and the rest as non-binary or self-identified. The age range distribution was 91 participants between 18–26, 26 participants between 27–39, 5 participants between 40–49, and 14 participants 50 or older. In terms of educational background, 61 participants held bachelor’s degrees, followed by 33 participants who had their under bachelor’s degrees, 23 participants who completed their master’s degrees, 8 participants who held honors qualifications, and the remaining participants had their Ph.D. or professional degrees.

3.4. Procedure

During the study, participants are initially presented with a study consent page. Upon agreeing to engage in this study, participants are then familiar with information related to SDGs, AI, and AI ethics, and the tasks involved in the study.

Subsequently, participants conduct the main tasks, which are presented in a randomized order to mitigate potential order effects. For each main task – corresponding to one of the 17 SDGs – participants

Table 3
AI ethical principles and their descriptions.

Ethical principle	Descriptions
Transparency	Explains how the AI systems make the decisions.
Justice & fairness	The degree to which predictions are unbiased and no systematic discrimination in AI systems.
Non-maleficence	No foreseeable or unintentional harm caused by AI systems.
Responsibility	AI systems act with ‘integrity’, with clear attribution of responsibility and legal liability.
Privacy	Sensitive information within AI systems is adequately protected.
Beneficence	AI systems should promote good, ensuring benefits for humanity, society, the environment, and the planet.
Freedom & autonomy	Humans retain the capacity for informed, autonomous decision-making within AI systems, free from manipulation.
Trust	The willingness of humans to rely on AI systems in situations characterized by uncertainty.
Sustainability	AI systems should create sustainable systems that process data sustainably and whose insights remain valid over time.
Dignity	AI systems should not diminish or destroy but respect, preserve, or increase human dignity.
Solidarity	The benefits from AI systems should not threaten social cohesion.

are asked to complete three subjective statements: (1) the perceived potential of AI in this SDG, (2) the perceived role of AI ethics when using AI in this SDG, (3) the perceived most important AI ethical principle when using AI in this SDG. For the first two statements, participants rate their agreement using a 5-point Likert scale. For the third statement, participants are required to determine the most important AI ethical principle from a list of eleven (for these principles and their descriptions, see Table 3). The main tasks of the study are designed to capture participants’ perceptions of AI and AI ethics across different SDGs and SDG groups.

After finishing 17 main tasks, participants are then asked to answer three questions about their personal characteristics on a demographic page: age, gender, and education level. During the conduction of the study, participants could not go back or change their past responses. This user study is approved by the Human Research Ethics Committee (HREC) of our university. Participation in the study is voluntary and uncompensated, with an estimated average completion time of 15 min.

4. Results

In this section, we first provide a detailed examination of how user perceptions vary across different SDGs (RQ1) in Section 4.1 and various SDG groups (RQ2) in Section 4.2, followed by a comprehensive overview of our research findings in Section 4.3. During data analysis, we perform a Friedman test and then follow it up with post-hoc



Fig. 1. Results for User Perceptions across SDGs and SDG Groups. In this figure, the (a), (b), and (c) refer to the perceived potential of AI, the perceived role of AI ethics, and the perceived most important AI ethical principle across SDGs, respectively. The (d), (e), and (f) refer to the perceived potential of AI, the perceived role of AI ethics, and the perceived most important AI ethical principle across SDG groups, respectively.

Conover tests (with a Bonferroni correction) to analyze differences in the perceived AI’s potential and the perceived role of AI ethics across different SDGs and SDG groups. Besides, we use a Rao-Scott chi-square test followed by post-hoc tests via pairwise Rao-Scott chi-square (with a Bonferroni correction) to analyze correlations between the perceived most important AI ethical principles and SDGs and SDG groups, respectively. Specifically, the Friedman test is an alternative to one-way ANOVA with repeated measures when the assumption of normally distributed data is violated [Goel et al. \(2022\)](#), [Friedman \(1937\)](#). Additionally, the Rao-Scott chi-square test is used when the dependent variable is categorical and independent observations are far from true [Scott \(2007\)](#), [Rao and Scott \(1981\)](#). [Fig. 1](#) shows the output plot for the user perceptions across distinct SDGs and SDG groups.

4.1. RQ1: How do different SDGs affect user perceptions of AI’s potential, the perceived role of AI ethics, and the perceived most important AI ethical principles?

To address RQ1 (related to H1), we quantify differences in the perceived potential of AI, the perceived role of AI ethics, and the perceived most important AI ethical principle from the seventeen SDGs. Subsequently, we apply a Friedman test followed by post-hoc Conover tests (with a Bonferroni correction) for the validation of H1.1 and H1.2. Besides, we conduct a Rao-Scott chi-square test followed by post-hoc pairwise Rao-Scott chi-square tests (with a Bonferroni correction) to validate H1.3. The results indicate significant differences in the perceived potential of AI, the perceived role of AI ethics across seventeen

SDGs, and reveal significant correlations between the perceived most important AI ethical principle and seventeen SDGs.

4.1.1. The perceived potential of AI across SDGs

The analysis of Friedman test results shows there are significant differences in the perceived potential of AI across various SDGs with $\chi^2(16) = 364.710, p < .001$. The post-hoc tests via Conover (with a Bonferroni correction) are conducted to find pairwise differences in the perceived potential of AI between the seventeen SDGs. There are significant differences being observed from Conover tests, as shown in [Fig. 2\(a\)](#). For example, the Conover tests find that the perceived potential of AI in SDG1 is significantly different from SDGs 3–9, and 11–17 (all $p < .001$, except SDG5: $p = .014$). Furthermore, Conover tests illuminate SDG2 to have significant differences in the perceived potential of AI from most other SDGs (SDGs 3–9, and 11–17; $p < .001$), with the exception of SDG5 ($p = .007$). Additionally, significant differences are found between SDG3 and several other SDGs (SDGs 5, 10, 16 ($p < .001$) and SDG15 ($p = .004$)). These outcomes suggest that the perceived potential of AI is affected by different SDGs, thus supporting H1.1.

4.1.2. The perceived role of AI ethics across SDGs

According to the Friedman test results, there are significant differences in the perceived role of AI ethics across various SDGs ($\chi^2(16) = 46.759, p < .001$). Then, post-hoc Conover tests (with a Bonferroni correction) are conducted to find differences in the perceived role of AI ethics among the seventeen SDGs. The Conover tests, based on

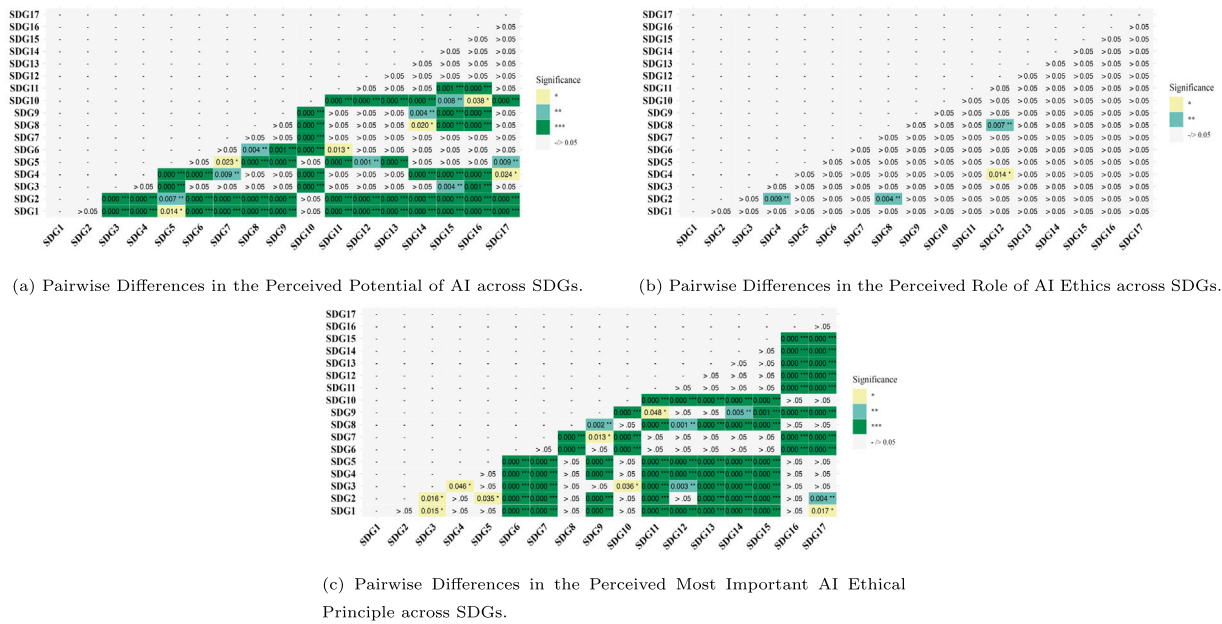


Fig. 2. Pairwise Differences in User Perceptions across SDGs. In this figure, the (a), (b), and (c) refer to the perceived potential of AI, the perceived role of AI ethics, and the perceived most important AI ethical principle across SDGs, respectively.

Fig. 2(b), indicate that the perceived role of AI ethics in SDG2 is significantly different from both SDG4 ($p = .009$) and SDG8 ($p = .004$). Moreover, Conover tests reveal SDG 12 to have significant variations in the perceived role of AI ethics from both SDG4 ($p = .014$) and SDG 8 ($p = .007$). The results imply that the perceived role of AI ethics is influenced by various SDGs, leading to the acceptance of H1.2.

4.1.3. The perceived most important AI ethical principles across SDGs

The Rao-Scott chi-square test indicates significant associations between the perceived most important AI ethical principle and SDG, $F(70.195, 9476.316) = 5.534, p < .001$. This suggests that the perceived most important AI ethical principle significantly vary across SDGs. Following significant associations, post-hoc pairwise Rao-Scott chi-square tests (with a Bonferroni correction) are performed to find differences in the perceived most important AI ethical principles among SDGs. The results of post-hoc pairwise Rao-Scott chi-square tests (Fig. 2(c)) reveal significant differences; for example, SDG 1 is significantly different from SDGs 6, 7, 9, and 11–15 ($p < .001$). Additionally, SDG10 exhibits significant differences in the perceived most important Ai ethical principles when comparing to SDGs 11–15 ($p < .001$) respectively. These results demonstrate that the perceived most important AI ethical principle is impacted by various SDGs, thereby confirming H1.3.

Overall, these findings demonstrate that the perceived potential of AI, the perceived of role of AI ethics, and the perceived most important AI ethical principles significantly vary across different SDGs. Thus, our H1.1, H1.2, and H1.3 are accepted.

4.2. RQ2: How do different SDG groups influence user perceptions of AI's potential, the perceived role of AI ethics, and the perceived most important AI ethical principles?

To answer RQ2 (related to H2), we assess variations in the perceived potential of AI, the perceived role of AI ethics, and the perceived most important AI ethical principle among three SDG groups (society, economy, and environment). Specifically, these groups are applied post hoc while participants are unaware of these groups during the study. We conduct a Friedman test followed by post-hoc Conover tests (with a Bonferroni correction) to validate both H2.1 and H2.2. Furthermore, to validate H2.3, we apply a Rao-Scott chi-square test followed by pairwise Rao-Scott chi-square tests (with a Bonferroni correction). Our

analysis reveals that the perceived potential of AI and the perceived most important AI ethical principles vary across different SDG groups. Nevertheless, no significant differences in the perceived role of AI ethics are found across different SDG groups.

4.2.1. The perceived potential of AI across SDG groups

The Friedman test results illustrate there are significant differences in the perceived potential of AI across three SDG groups ($\chi^2(2) = 13.546, p = .001$). The post-hoc Conover tests (with a Bonferroni correction) are further applied to find pairwise differences in the perceived potential of AI when comparing three SDG groups. According to Fig. 3(a), the Conover tests indicate that the perceived potential of AI in economy is significantly different from both that in society ($p = .002$) and environment ($p = .008$). These findings imply that the perceived potential of AI is different for various SDG groups, thereby leading to the acceptance of H2.1.

4.2.2. The perceived role of AI ethics across SDG groups

A Friedman test is conducted to explore variations in the perceived role of AI ethics across three SDG groups. The analysis reveals no statistically significant difference in the perceived role of AI ethics among SDG groups, leading to the rejection of H2.2.

4.2.3. The perceived most important AI ethical principles across SDG groups

A Rao-Scott chi-square test is applied to investigate the associations between the perceived most important AI ethical principles and SDG groups. For this test, we weight society, economy, and environment as $\frac{1}{9}, \frac{1}{5},$ and $\frac{1}{3}$ to account for unequal SDG distributions (9, 5, and 3 SDGs per group, respectively). No weighting is necessary in prior Friedman analyses. The analysis of Rao-Scott chi-square test results shows significant associations between the perceived most important AI ethical principle and SDG groups, $F(15.237, 2057.056) = 16.022, p < .001$. This suggests that the perceived most important AI ethical principle significantly vary across SDG groups. Furthermore, post-hoc pairwise Rao-Scott chi-square tests are employed to find differences in the perceived most important AI ethical principle across SDG groups. Regarding Fig. 3(b), the perceived most important AI ethical principles in society differs significantly from that in economy ($p = .003$) and environment ($p < .001$), respectively. Moreover, there are significant

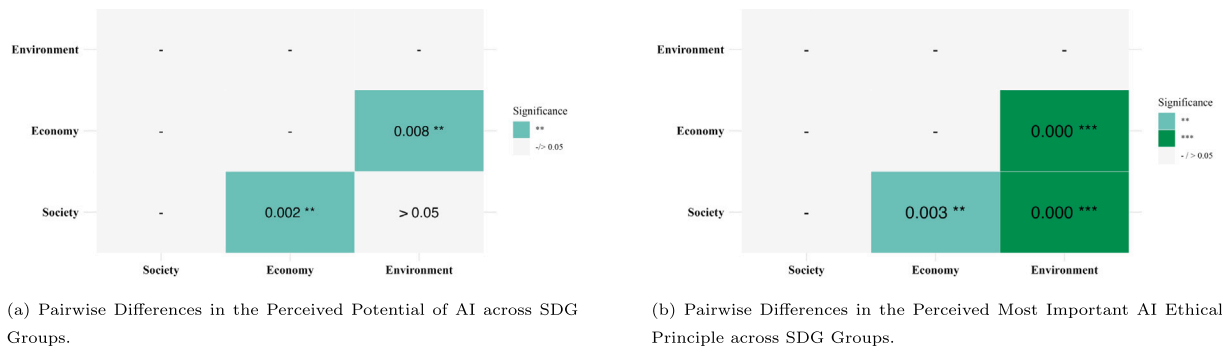


Fig. 3. Pairwise Differences in User Perceptions across SDG Groups. In this figure, the (a) and (b) refer to the perceived potential of AI and the perceived most important AI ethical principle across SDG groups, respectively.

Table 4
Minimum Detectable Effect Size (MDES) and Observed Effect Size (OES) with 95% Confidence Intervals for User Perceptions across SDGs and SDG Groups.

	MDES	OES with Intervals
H1.1	0.065	0.168, [0.145, 0.200]
H1.2	0.065	0.021, [0.017, 0.037]
H1.3	0.123	0.130, [0.111, 0.148]
H2.1	0.065	0.050, [0.019, 0.101]
H2.2	0.065	0.010, [0.001, 0.041]
H2.3	0.090	0.230, [0.189, 0.271]

differences between economy and environment ($p < .001$). These outcomes imply that the perceived most important AI ethical principle varies significantly across SDG groups, leading to accept H2.3. In summary, these results indicate that the perceived potential of AI and the perceived most important AI ethical principles are influenced by SDG groups. As a result, our H2.1 and H2.3 are accepted. However, there are no significant differences in the perceived role of AI ethics among SDG groups, leading to the rejection of H2.2.

4.3. Findings

In this study, we conduct a human-based experimental study introducing the SDGs scenario to examine the impact of seventeen SDGs on user perceptions—the perceived potential of AI, the perceived role of AI ethics, and the perceived most important AI ethical principles. After primary analysis of SDGs’ influence on user perceptions, we have a post-hoc analysis about the effects of three SDG groups (society, economy, and environment) on user perceptions. Table 4 summarizes the magnitude of the minimum detectable effect size (MDES) at 80% power and the observed effect size with confidence intervals for user perceptions across SDGs and SDG groups. Our findings are discussed as follows:

User perceptions across SDGs. Our results show that there are significant variations in the perceived potential of AI, the perceived role of AI ethics, and the perceived most important AI ethical principle across different SDGs. Beyond statistical significance, the observed effect sizes for H1.1 (0.168, 95% CI [0.145, 0.200]) and H1.3 (0.130, 95% CI [0.111, 0.148]) exceeded their MDES (0.065 and 0.123) respectively based on Table 4, indicating a robust and practically meaningful difference in perceived AI potential and perceived AI ethical principles across SDGs. For H1.2, the observed effect size (0.021, 95% CI [0.017, 0.037]) falls well below the MDES (0.065). This suggests that although the variation in the perceived AI role across is statistically different, its practical significance is limited. Hence, these findings affirm H1.1 and H1.3, while highlighting the limited practical impact for H1.2.

User perceptions across SDG groups. Results indicate that the perceived potential of AI and the perceived most important AI ethical principle significantly differ in three SDG groups. Based on Table 4, H2.1 shows a small observed effect size (0.050, 95% CI [0.019, 0.101]) below the MDES (0.065), indicating a limited practical significance. However, the observed effect size (0.230, 95% CI [0.189, 0.271]) for H2.3 exceeds the MDES (0.090), demonstrating a robust and practically significant difference. In contrast, no significant differences in the perceived role of AI ethics are observed across SDG groups, thereby leading to the rejection of H2.2 and suggesting that participants regard ethical considerations as universally important rather than domain-specific. One possible explanation is that the role of AI ethics is perceived as foundational requirements for AI in any societal, economic, or environment context. Theoretically, this finding indicates that AI ethics are regarded as essential (see Fig. 1(e)), even in the absence of significant differences across SDG groups.

From a theoretical standpoint, this result may indicate a shared mental model of AI ethics that transcends contextual boundaries, reinforcing the argument for developing cross-cutting ethical governance frameworks. Practically, it implies that policymakers and designers might prioritize universally applicable ethical standards while allowing for context-specific adaptations where necessary. Future research could investigate whether this universality persists in more concrete, domain-specific scenarios or when participants are exposed to real-world AI use cases within different SDGs.

In summary, our study first confirms that SDGs significantly and meaningfully influence user perceptions, encompassing the perceived potential of AI, and the perceived most important AI ethical principle. Additionally, our study demonstrates that while SDG groups affect the perceived potential of AI with a limited practical significance, and the perceived most important AI ethical principle with a robust significance; however, their impacts on the perceived role of AI ethics are less evident.

5. Managerial insights

In this section, we delineate the theoretical and practical implications of these results. Subsequently, we provide a discussion of the limitations and future research directions of our study.

5.1. Implications

Our empirical findings are crucial for advancing the development of human-centric AI within the context of the SDGs. In some cases, AI technologies pose significant challenges rather than serving as solutions (Kusner and Loftus, 2020). As such, AI may serve as either an enabler or an inhibitor in achieving the 17 SDGs. To effectively harness AI’s potential while mitigating its risks, it must be integrated with AI ethical frameworks and human-centric perspectives. Regarding AI ethics, it concerns not only the moral behavior of AI itself, but also

the moral behavior of those who design, develop, and deploy AI. With respect to human perspectives, AI must be applied in accordance with human rights. The integration of AI, AI ethics, and human values is essential for the realization of human-centric AI. Ultimately, human-centric AI can perform operations and tasks aimed at improving the welfare of all, caring for the planet, societies, and individuals. Specifically, this perspective aligns with SDG 11, building communities that integrate AI systems to enhance societal wellbeing. Such cities and communities leveraging human-centered AI development as a means of sharing knowledge and coordinating actions will make it possible to construct new social models and interactions, thereby strengthening the human fabric (Pastor-Escuredo et al., 2022).

Additionally, our outcomes hold valuable relevance for policymakers, developers, and scholars who are tasked with the field of ethical AI systems in the context of SDG. Grounded in societal values, the results inform the design and deployment of AI systems aligned with SDGs. For example, policymakers could leverage these insights to craft and complete regulations that address public concerns and prioritize context-specific ethical principles across SDGs. Moreover, developers can integrate ethical principles that people deem most important when designing and implementing AI systems in SDG-related applications. Furthermore, for scholars, they can build on these initial findings to explore broader ethical questions, such as whether AI can help humans develop a more holistic understanding of sustainable development needs, as suggested by Pastor-Escuredo et al. (2021). Collectively, these findings underscore the importance of integrating user perceptions of AI and AI ethics into SDG initiatives.

5.2. Limitations and future directions

Though our study contributes novel findings about the user perceptions of AI, AI ethics, and AI ethical principle across SDG and SDG groups, specific inherent limitations in the experimental design would be acknowledged. One such limitation is that the experiment presents all SDGs to users in text to elicit user perceptions. Due to the wide scope and number of SDGs, this approach is a practical means to address our research questions. However, real-world AI systems or use cases within SDG contexts are often much complex, and user perceptions related to AI systems are highly context-dependent (Kieslich et al., 2022). Also, the context-comparing study by Kieslich et al. (2021) has concluded that users' threat perceptions regarding AI are largely domain-specific. As such, perceptions may vary significantly depending on the specific characteristics of AI systems associated with individual SDGs. Future research should consider simulating concrete AI use cases aligned with specific SDGs to better capture context-sensitive user perceptions (Kieslich et al., 2022, 2021).

Another limitation of this study lies in the participant sample, which is drawn primarily from a limited number of countries, such as Australia, China, and Austria. Given that SDGs are inherently global in scope, this narrow participant pool constrains our analysis to capture culturally nuanced perspectives on AI's societal and ethical implications. Moreover, countries in different cultures adopt various levels and applications of AI technologies and follow various national strategies for the development of AI and AI ethics because of their tradition and religion (Pastor-Escuredo and Vinueza, 2020; Kieslich et al., 2022). As a result, user perceptions of AI and AI ethics may vary significantly across countries. Thereby, incorporating participants from a more diverse range of countries would enhance a more comprehensive and culturally informed analysis of user perceptions of AI and AI ethics in the future (Kelley et al., 2021).

Besides, some observed effect sizes from our results are smaller than the MDES determined for 80% power. It indicates that our study is underpowered to detect very small effects reliably. Future work could increase sample size and optimize design to detect tiny differences in user perceptions across SDGs and SDG groups (Funder and Ozer, 2019). At last, the limitation of our study is that our metrics rely solely on one

subjective metric, which may introduce self-reported biases, particularly given the use of absolute phrasing in some questions (e.g., end poverty) that participants may perceive as unrealistic. It is possible that there is a difference between a subjective (attentive) and an objective (non-attentive) result due to a distinction between a 'mindful' (attentive) and a 'mindless' (non-attentive) state (Langer et al., 1978). Therefore, future study should employ a combination of subjective and objective metrics to measure users' perceptions (Sanneman and Shah, 2022).

6. Conclusion

With the progress and advances of AI, it is possible that AI can enable or inhibit the delivery of SDGs. Let us say that AI-driven automation enables SDG 12 by enhancing productivity, yet hinders SDG 10 by limiting labor market opportunities for low-skilled workers. AI ethics, concerned with addressing the ethical issues of AI, is essential to ensuring that AI systems operate ethically. Given that both ethical AI and the SDGs aim to be human-centric and socially beneficial, incorporating user perceptions is critical. Therefore, in this paper, we conduct a user study within the context of the SDGs to examine how individual SDGs and SDG groups influence user perceptions, specifically regarding the perceived potential of AI, the perceived role of AI ethics, and the perceived most important AI ethical principle. Our findings indicate significant differences in the perceived potential of AI and ethical priorities depending on the SDGs and SDG groups, while the perceived role of AI ethics remains consistent across goal groups. These insights highlight the importance of incorporating user perspectives into discussions of AI and AI ethics in the SDG context. Advancing the SDGs through ethical AI requires a commitment to human values, ensuring progress that is ethical, inclusive, and aligned with societal needs.

In addition, this study has several limitations, including the absence of concrete AI application scenarios, a sample drawn from a limited number of countries, limited power to detect very small effects, and potential self-report biases. Future research should address these limitations by examining specific AI use cases within particular SDG contexts and incorporating cross-national samples, along with a combination of subjective and objective evaluation metrics.

CRedit authorship contribution statement

Bo Wang: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Data curation, Conceptualization. **Jianlong Zhou:** Writing – review & editing, Visualization, Supervision, Project administration, Methodology, Conceptualization. **Fang Chen:** Writing – review & editing, Supervision. **Heimo Müller:** Writing – review & editing, Supervision. **Andreas Holzinger:** Writing – review & editing, Supervision, Conceptualization.

Declaration of AI

The authors state that they only used tools such as Grammarly, Writefull, and DeepL for language and spelling improvement, but confirm that they did not use any generative AI for content creation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Parts of this work have been funded by the Austrian Science Fund (FWF), Project: P-32554 explainable Artificial Intelligence. The ethical approval for our research is granted by the HREC of the University of Technology Sydney with the number ETH22-7616. We thank the participants who took part in our studies.

References

- Acemoglu, D., Restrepo, P., 2018. Artificial intelligence, automation, and work. In: *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press, pp. 197–236. <http://dx.doi.org/10.3386/w24196>.
- Agency, I.E., 2017. Digitalization & energy. URL <https://iea.blob.core.windows.net/assets/b1e6600c-4e40-4d9c-809d-1d1724c763d5/DigitalizationandEnergy3.pdf>. (Accessed 20 March 2025).
- Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., Mané, D., 2016. Concrete problems in AI safety. <http://dx.doi.org/10.48550/arXiv.1606.06565>, arXiv preprint [arXiv:1606.06565](https://arxiv.org/abs/1606.06565).
- Angerschmid, A., Zhou, J., Theuermann, K., Chen, F., Holzinger, A., 2022. Fairness and explanation in AI-informed decision making. *Mach. Learn. Knowl. Extr.* 4 (2), 556–579. <http://dx.doi.org/10.3390/make4020026>.
- Astobiza, A.M., Toboso, M., Aparicio, M., López, D., 2021. AI Ethics for Sustainable Development Goals. *IEEE Technol. Soc. Mag.* 40 (2), 66–71. <http://dx.doi.org/10.1109/MTS.2021.3056294>, URL <https://ieeexplore.ieee.org/document/9445792/?number=9445792>. Conference Name: IEEE Technology and Society Magazine.
- Banks, S., Formosa, P., 2023. The ethical implications of artificial intelligence (AI) for meaningful work. *J. Bus. Ethics* 185 (4), 725–740. <http://dx.doi.org/10.1007/s10551-023-05339-7>.
- Berendt, B., 2019. AI for the common good?! pitfalls, challenges, and ethics pen-testing. *Paladyn J. Behav. Robot.* 10 (1), 44–65. <http://dx.doi.org/10.48550/arXiv.1810.12847>.
- Bolkubasi, T., Chang, K.-W., Zou, J.Y., Saligrama, V., Kalai, A.T., 2016. Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *Adv. Neural Inf. Process. Syst.* 29, <http://dx.doi.org/10.48550/arXiv.1607.06520>.
- Bourhis, P., Demartini, G., Elbassouni, S., Hoareau, E., Rao, H.R., 2019. Ethical challenges in the future of work. *Bull. Tech. Comm. Data Eng.* URL <https://hal.science/hal-03094585/document>.
- Bundy, A., 2017. Preparing for the future of artificial intelligence. <http://dx.doi.org/10.1007/s00146-016-0685-0>.
- Courtland, R., 2018. Bias detectives: the researchers striving to make algorithms fair. *Nature* 558 (7710), 357–360. <http://dx.doi.org/10.1038/d41586-018-05469-3>.
- Cowls, J., Tsamados, A., Taddeo, M., Floridi, L., 2021. A definition, benchmark and database of AI for social good initiatives. *Nat. Mach. Intell.* 3 (2), 111–115. <http://dx.doi.org/10.1038/s42256-021-00296-0>.
- Floridi, L., Cowls, J., Beltramini, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., et al., 2018. AI4people—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. *Minds Mach.* 28, 689–707. <http://dx.doi.org/10.1007/s11023-018-9482-5>.
- Friedman, M., 1937. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. Statist. Assoc.* 32 (200), 675–701. <http://dx.doi.org/10.2307/2279372>.
- Funder, D.C., Ozer, D.J., 2019. Evaluating effect size in psychological research: Sense and nonsense. *Adv. Methods Pr. Psychol. Sci.* 2 (2), 156–168. <http://dx.doi.org/10.1177/2515245920979282>.
- Fuso Nerini, F., Slob, A., Ericsson Engström, R., Trutnevte, E., 2019. A research and innovation agenda for zero-emission European cities. *Sustainability* 11 (6), 1692. <http://dx.doi.org/10.3390/su11061692>.
- Glavič, P., 2021. Evolution and current challenges of sustainable consumption and production. *Sustainability* 13 (16), 9379. <http://dx.doi.org/10.3390/su13169379>.
- Goel, K., Sindhgatta, R., Kalra, S., Goel, R., Mutreja, P., 2022. The effect of machine learning explanations on user trust for automated diagnosis of COVID-19. *Comput. Biol. Med.* 146, 105587. <http://dx.doi.org/10.1016/j.compbiomed.2022.105587>.
- Gupta, A., Lanteigne, C., Heath, V., Ganapini, M.B., Galinkin, E., Cohen, A., De Gasperis, T., Akif, M., Butalid, R., 2020. The state of AI ethics report (June 2020). <http://dx.doi.org/10.48550/arXiv.2006.14662>, arXiv preprint [arXiv:2006.14662](https://arxiv.org/abs/2006.14662).
- Hoffman, R.R., Mueller, S.T., Klein, G., Litman, J., 2018. Metrics for explainable AI: Challenges and prospects. <http://dx.doi.org/10.48550/arXiv.1812.04608>, arXiv preprint [arXiv:1812.04608](https://arxiv.org/abs/1812.04608).
- Holzinger, A., Zatloukal, K., Müller, H., 2025. Is human oversight to AI systems still possible? *New Biotechnol.* 85, 59–62. <http://dx.doi.org/10.1016/j.nbt.2024.12.003>.
- Huang, C., Zhang, Z., Mao, B., Yao, X., 2022. An overview of artificial intelligence ethics. *IEEE Trans. Artif. Intell.* 4 (4), 799–819. <http://dx.doi.org/10.1109/TAI.2022.3194503>.
- Jobin, A., Ienca, M., Vayena, E., 2019. Artificial Intelligence: the global landscape of ethics guidelines. *Nat. Mach. Intell.* 1 (9), 389–399. <http://dx.doi.org/10.1038/s42256-019-0088-2>, URL <http://arxiv.org/abs/1906.11668>. arXiv:1906.11668 [cs].
- Kelley, P.G., Yang, Y., Heldreth, C., Moessner, C., Sedley, A., Kramm, A., Newman, D.T., Woodruff, A., 2021. Exciting, useful, worrying, futuristic: Public perception of artificial intelligence in 8 countries. In: *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*. pp. 627–637. <http://dx.doi.org/10.1145/3461702.3462605>.
- Keramitsoglou, I., Cartalis, C., Kiranoudis, C.T., 2006. Automatic identification of oil spills on satellite images. *Environ. Model. Softw.* 21 (5), 640–652. <http://dx.doi.org/10.1016/j.envsoft.2004.11.010>.
- Kieslich, K., Keller, B., Starke, C., 2022. Artificial intelligence ethics by design. Evaluating public perception on the importance of ethical design principles of artificial intelligence. *Big Data Soc.* 9 (1), 20539517221092956. <http://dx.doi.org/10.1177/20539517221092956>.
- Kieslich, K., Lünich, M., Marcinkowski, F., 2021. The threats of artificial intelligence scale (TAI) development, measurement and test over three application domains. *Int. J. Soc. Robot.* 13 (7), 1563–1577. <http://dx.doi.org/10.1007/s12369-020-00734-w>.
- Kusner, M.J., Loftus, J.R., 2020. The long road to fairer algorithms. *Nature* 578 (7793), 34–36. <http://dx.doi.org/10.1038/d41586-020-00274-3>.
- Langer, E.J., Blank, A., Chanowitz, B., 1978. The mindlessness of ostensibly thoughtful action: The role of "placebic" information in interpersonal interaction. *J. Pers. Soc. Psychol.* 36 (6), 635. <http://dx.doi.org/10.1037/0022-3514.36.6.635>.
- Lo Piano, S., 2020. Ethical principles in machine learning and artificial intelligence: cases from the field and possible ways forward. *Humanit. Soc. Sci. Commun.* 7 (1), 1–7. <http://dx.doi.org/10.1057/s41599-020-0501-9>.
- Mohamadi, A., Heidarzadi, Z., Nourollahi, H., 2016. Assessing the desertification trend using neural network classification and object-oriented techniques (case study: Changouleh watershed-iland province of Iran). *J. Fac. For. Istanbul Univ.* 66 (2), 683–690. <http://dx.doi.org/10.17099/jffiu.75819>.
- Nagano, A., 2018. Economic growth and automation risks in developing countries due to the transition toward digital modernity. In: *Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance*. pp. 42–50. <http://dx.doi.org/10.1145/3209415.3209442>.
- Pastor-Escuredo, D., Giacomelli, G., Lumberas, J., Garbajosa, J., 2021. Ethical and sustainable Future OF Work. *Rev. Diecisiete: Investigación Interdisciplinaria Para Los Objet. Desarro. Sosten.* 04, 183–191. http://dx.doi.org/10.36852/2695-4427_2021_04.11.
- Pastor-Escuredo, D., Treleven, P., Vinuesa, R., 2022. An Ethical Framework for Artificial Intelligence and Sustainable Cities. *AI* 3 (4), 961–974. <http://dx.doi.org/10.3390/ai3040057>, URL <https://www.mdpi.com/2673-2688/3/4/57>. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.
- Pastor-Escuredo, D., Vinuesa, R., 2020. Towards an ethical framework in the complex digital era. <http://dx.doi.org/10.48550/arXiv.2010.10028>, arXiv preprint [arXiv:2010.10028](https://arxiv.org/abs/2010.10028).
- Rao, J.N., Scott, A.J., 1981. The analysis of categorical data from complex sample surveys: chi-squared tests for goodness of fit and independence in two-way tables. *J. Amer. Statist. Assoc.* 76 (374), 221–230. <http://dx.doi.org/10.2307/2287815>.
- Ribeiro, M.T., Singh, S., Guestrin, C., 2016. "Why should I trust you?" explaining the predictions of any classifier. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. pp. 1135–1144. <http://dx.doi.org/10.48550/arXiv.1602.04938>.
- Rietveld, T., Van Hout, R., 2007. Analysis of variance for repeated measures designs with word materials as a nested random or fixed factor. *Behav. Res. Methods* 39, 735–747. <http://dx.doi.org/10.3758/BF03192964>.
- Ryan, M., Antoniou, J., Brooks, L., Jiya, T., Macnish, K., Stahl, B., 2020. The ethical balance of using smart information systems for promoting the united nations' sustainable development goals. *Sustainability* 12 (12), 4826. <http://dx.doi.org/10.3390/su12124826>.
- Sadowski, M., Powell-Tuck, R., 2017. AI & the sustainable development goals: The state of play. URL <https://www.sustainability.com/globalassets/sustainability.com/thinking/pdfs/2030vision-stateofplay.pdf>. (Accessed 12 August 2025).
- Sanneman, L., Shah, J.A., 2022. The situation awareness framework for explainable AI (SAFE-AI) and human factors considerations for XAI systems. *Int. J. Human-Computer Interact.* 38 (18–20), 1772–1788. <http://dx.doi.org/10.1080/10447318.2022.2081282>.
- Scott, A., 2007. Rao-scott corrections and their impact. In: *Proceedings of the 2007 Joint Statistical Meetings, Salt Lake City, Utah*. pp. 3514–3518, URL <https://api.semanticscholar.org/CorpusID:195854296>.
- Siau, K., Wang, W., 2020. Artificial intelligence (AI) ethics: ethics of AI and ethical AI. *J. Database Manag. (JDM)* 31 (2), 74–87. <http://dx.doi.org/10.4018/JDM.2020040105>.
- Society, I., 2017. Artificial Intelligence & Machine Learning: Policy Paper. URL <https://www.internetsociety.org/resources/doc/2017/artificial-intelligence-and-machine-learning-policy-paper/>. (Accessed 18 March 2025).
- Stockholm University, 2017. Stockholm resilience centre's (SRC) contribution to the 2016 Swedish 2030 agenda HLPF report. URL <https://www.stockholmresilience.org/download/18.2561f5bf15a1a341a523695/1488272270868/SRCs+2016+Swedish+2030+Agenda+HLPF+report+Final.pdf>. (Accessed 25 March 2025).
- The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, 2017. Ethically aligned design - version II. URL https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.2.pdf. (Accessed 25 March 2025).
- Tomašev, N., Cornebise, J., Hutter, F., Mohamed, S., Picciariello, A., Connelly, B., Belgrave, D.C., Ezer, D., Haert, F.C.v.d., Mugisha, F., et al., 2020. AI for social good: unlocking the opportunity for positive impact. *Nat. Commun.* 11 (1), 2468. <http://dx.doi.org/10.1038/s41467-020-15871-z>.
- Truby, J., 2018. Decarbonizing Bitcoin: Law and policy choices for reducing the energy consumption of blockchain technologies and digital currencies. *Energy Res. Soc. Sci.* 44, 399–410. <http://dx.doi.org/10.1016/j.erss.2018.06.009>.
- United Nations Department of Economic and Social Affairs, 2015. Sustainable development. URL <https://sdgs.un.org/>. (Accessed 28 March 2025).

- Varriale, V., Cammarano, A., Michelino, F., Caputo, M., 2024. The role of digital technologies in production systems for achieving sustainable development goals. *Sustain. Prod. Consum.* 47, 87–194. <http://dx.doi.org/10.1016/j.spc.2024.03.035>.
- Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S.D., Tegmark, M., Fuso Nerini, F., 2020. The role of artificial intelligence in achieving the Sustainable Development Goals. *Nat. Commun.* 11 (1), 233. <http://dx.doi.org/10.1038/s41467-019-14108-y>, URL <https://www.nature.com/articles/s41467-019-14108-y>. Publisher: Nature Publishing Group.
- Yeh, S.C., Wu, A.W., Yu, H.C., Wu, H.C., Kuo, Y.P., Chen, P.X., 2021. Public perception of artificial intelligence and its connections to the sustainable development goals. *Sustainability* 13 (16), 9165. <http://dx.doi.org/10.3390/su13169165>.
- Zheng, L., Chiang, W.L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y., Lin, Z., Li, Z., Li, D., Xing, E., et al., 2023. Judging llm-as-a-judge with mt-bench and chatbot arena. *Adv. Neural Inf. Process. Syst.* 36, 46595–46623. <http://dx.doi.org/10.48550/arXiv.2306.05685>.
- Zhou, J., Chen, F., 2018. *Human and Machine Learning*. Springer, <http://dx.doi.org/10.1007/978-3-319-90403-0>.
- Zhou, J., Chen, F., 2023. AI ethics: from principles to practice. *AI & SOCIETY* 38 (6), 2693–2703. <http://dx.doi.org/10.1007/s00146-022-01602-z>.
- Zhou, J., Chen, F., Berry, A., Reed, M., Zhang, S., Savage, S., 2020. A survey on ethical principles of AI and implementations. In: 2020 IEEE Symposium Series on Computational Intelligence. SSCI, IEEE, pp. 3010–3017. <http://dx.doi.org/10.1109/SSCI47803.2020.9308437>.
- Zhou, J., Müller, H., Holzinger, A., Chen, F., 2024. Ethical ChatGPT: Concerns, challenges, and commandments. *Electronics* 13 (17), <http://dx.doi.org/10.3390/electronics13173417>.