Technological Forecasting & Social Change 220 (2025) 124304

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

Mapping public perception of artificial intelligence: Expectations, risk–benefit tradeoffs, and value as determinants for societal acceptance

Philipp Brauner *, Felix Glawe, Gian Luca Liehner, Luisa Vervier, Martina Ziefle

Communication Science RWTH Aachen University, 52056, Aachen, Germany

ARTICLE INFO

Dataset link: https://osf.io/6q3eu/

Keywords:
AI ethics
AI alignment
Technology acceptance
Risk perception
Technology assessment
Mental models
Artificial intelligence

ABSTRACT

Public opinion on artificial intelligence (AI) plays a pivotal role in shaping trust and AI alignment, ethical adoption, and the development equitable policy frameworks. This study investigates expectations, risk–benefit tradeoffs, and value assessments as determinants of societal acceptance of AI. Using a nationally representative sample (N = 1100) from Germany, we examined mental models of AI and potential biases. Participants evaluated 71 AI-related scenarios across domains such as autonomous driving, medical care, art, politics, warfare, and societal divides, assessing their expected likelihood, perceived risks, benefits, and overall value. We present ranked evaluations alongside visual mappings illustrating the risk–benefit tradeoffs. Our findings suggest that while many scenarios were considered likely, they were often associated with high risks, limited benefits, and low overall value. Regression analyses revealed that 96.5% (r^2 = 0.965) of the variance in value judgments was explained by risks (β = −0.490) and, more strongly, benefits (β = +0.672), with no significant relationship to expected likelihood. Demographics and personality traits, including age, gender, and AI readiness, influenced perceptions, highlighting the need for targeted AI literacy initiatives. These findings offer actionable insights for researchers, developers, and policymakers, highlighting the need to communicate tangible benefits and address public concerns to foster responsible and inclusive AI adoption. Future research should explore cross-cultural differences and longitudinal changes in public perception to inform global AI governance.

1. Introduction

Rapid advancements in Artificial Intelligence (AI) and Deep Learning (DL), particularly in large language models (LLMs)s, have generated widespread interest and concern across multiple domains. Although the roots of AI date back many decades (McCarthy et al., 2006; Hopfield, 1982; Rumelhart et al., 1986), recent progress has been driven by increased computational power, the growing availability of digital training data (Deng et al., 2009), more sophisticated algorithms (Lecun et al., 2015), and a significant rise in funding (Statista, 2022). AI is now increasingly integrated into sectors ranging from education (Chen et al., 2020) and healthcare (Amunts et al., 2023) to journalism (Diakopoulos, 2019), forestry and agriculture (Holzinger et al., 2024), and manufacturing (Brauner et al., 2022).

AI promises numerous benefits, including enhanced efficiency, convenience, and innovation (Bouschery et al., 2023). However, it also raises concerns about privacy, labor displacement (Acemoglu and Restrepo, 2017), and complex ethical dilemmas affecting individuals, organizations, and society at large (Awad et al., 2018). Consequently, expectations regarding AI are polarized. While some view it as a transformative force capable of improving various aspects of life (Brynjolfsson and McAfee, 2014; Makridakis, 2017; Bouschery et al., 2023), others emphasize its ethical challenges and societal risks (Cath, 2018; Bostrom, 2003; Crawford, 2021).

Researchers have long recognized that computing technologies and algorithms are not inherently value-neutral; rather, they often reflect and reinforce underlying social values and biases (Forsythe, 1993; Friedman and Nissenbaum, 1996; Nissenbaum, 2001). These embedded values can influence decisions and outcomes, potentially perpetuating inequalities or exacerbating existing social disparities (Budish, 2021; Garcia, 2024; Mittelstadt et al., 2016). Critically examining public perceptions of AI and its broader social implications is thus essential. A central issue in this context is the so-called alignment problem: the challenge of ensuring that AI systems align with human values such as fairness, transparency, and accountability (Gabriel, 2020; Hristova et al., 2024).

As concerns about AI's ethical and societal impacts grow, researchers and policymakers increasingly emphasize the need for deliberate design choices and anticipatory governance to mitigate

* Corresponding author.
E-mail address: brauner@comm.rwth-aachen.de (P. Brauner).

https://doi.org/10.1016/j.techfore.2025.124304
Received 29 January 2025; Received in revised form 30 June 2025; Accepted 22 July 2025
Available online 16 August 2025

negative consequences (Gursoy and Kakadiaris, 2023). Understanding how the public perceives AI—particularly the trade-offs between perceived risks and benefits—is crucial, as these perceptions shape policy decisions, innovation trajectories, and public acceptance (Sadek et al., 2024).

While prior research has typically examined either specific AI applications or general attitudes towards AI, this study adopts an integrative perspective. We investigate how individuals evaluate a wide range of projected AI scenarios. Using an online survey of a nationally representative sample in Germany, participants assessed many AI-related scenarios anticipated over the next decade in terms of expected likelihood, perceived risks and benefits, as well as overall attributed value.

Our analysis proceeds on two levels. At the individual level, we explore how demographic characteristics and personality traits influence AI perceptions. At the scenario level, we examine how assessments of risk and benefit inform value judgments. We also employ risk–benefit maps to visualize where public expectations converge or diverge. Results show that overall perceived value is primarily shaped by perceived benefits, with perceived risks playing a significant, but secondary role. Demographic variables further influence these evaluations, though increasing AI literacy may help reduce disparities.

By identifying key patterns of alignment and divergence in public perception, this study offers new insights into potential tensions in societal acceptance of AI. It provides practical guidance for researchers, developers, policymakers, and educators aiming to align AI development with public values and improve communication and engagement strategies.

The remainder of the article is structured as follows. Section 2 reviews literature on public perceptions of AI and relevant findings from technology and risk perception research. Section 3 outlines the survey methodology and introduces the micro-scenario approach. Section 4 presents the findings, beginning with scenario evaluations and visual mappings, followed by an analysis of individual-level differences. Section 5 discusses the broader implications and limitations of the study. Section 6 concludes with policy recommendations and suggestions for future research.

2. Related work

Liu (2021) categorizes social science research on AI into three perspectives: (1) the scientific perspective, which analyzes AI as a research domain; (2) the technical perspective, which considers AI as a meta-technology embedded across various technological systems and artifacts; and (3) the cultural perspective, which examines AI as a socio-cultural phenomenon shaped by and shaping individual, social, economic, and political contexts. This study aligns with the cultural perspective, as it explores public perceptions of AI technologies within broader societal frameworks.

This section reviews current literature on public perceptions of AI. It begins by surveying the general landscape of research on AI perception. It then considers how perceptions of risk and benefit provide a conceptual foundation for analyzing public responses to emerging technologies, including AI. The following subsection explores how the application context influences public attitudes towards AI. Finally, the section reviews how individual-level differences—such as demographics and personality traits—shape these perceptions. The section concludes by identifying gaps in existing research and introducing the research questions guiding this study.

2.1. General AI perception

Recent studies on public perceptions of AI reveal a diverse and fragmented landscape, both in terms of methodological approaches and empirical findings. Fast and Horvitz (2017) conducted a longitudinal media analysis of The New York Times and identified a notable increase in public interest in AI since 2009. This shift was marked by a mix of optimism and concern, with generally more favorable than unfavorable coverage. However, recent trends indicate a growing emphasis on risks such as loss of control and ethical dilemmas, especially in contrast to earlier optimism, particularly surrounding AI's applications in healthcare. The study also noted that public expectations of AI are often inflated or unrealistic, influenced by optimistic portrayals in news and entertainment media.

Building on this line of inquiry, Sanguinetti and Palomo (2024) introduced an "AI Anxiety Index" to assess how major newspapers frame public apprehension about AI. Their analysis of headlines before and after the release of ChatGPT showed a significant uptick in AI-related coverage, accompanied by increasingly negative sentiment. Notably, regional media played a disproportionate role in amplifying AI-related anxiety by portraying AI as an autonomous and opaque force beyond human control.

Despite the rising media attention, public understanding of AI remains limited. Survey data from Ipsos indicates that many individuals struggle to distinguish between AI and related concepts such as robotics, automation, and machine learning (Ipsos, 2022). Similarly, Pew Research found that only a minority of Americans could accurately identify AI technologies in everyday contexts, underscoring a widespread lack of clarity about AI's capabilities and boundaries (Pew Research Center, 2023). This knowledge gap may hinder informed public discourse and exacerbate misconceptions about AI's societal and ethical implications. The Alan Turing Institute further highlighted the contextual variability of public understanding, noting that concerns often focus on employment and security-related automation, and are influenced by education level and media framing (Anon, 2023).

Public discourse around AI often combines vague fears with overly optimistic expectations, particularly regarding artificial general intelligence (AGI), which remains a largely speculative concept (Jungherr, 2023). Cave et al. (2019) explored AI narratives in the UK and identified eight dominant themes—four positive and four negative. Their study found that many participants viewed AI through a lens of anxiety and uncertainty, with only two narratives suggesting that the benefits outweigh the risks. A substantial proportion of respondents expressed feelings of disempowerment, viewing AI development as driven by governments and corporations. Approximately half could provide plausible definitions of AI, while 25% associated the concept primarily with robots.

In a prior study, we examined laypersons' expectations (perceived likelihood) and sentiment (positive or negative outlook) regarding AI-related scenarios, using a younger, convenience-based sample (Brauner et al., 2023). The findings revealed notable variation in public perception. In particular, scenarios involving cybersecurity threats were considered both highly likely and least desirable, indicating heightened concern in this domain.

Growing awareness of biased algorithms and discriminatory outcomes has further intensified public concern about the ethical use of AI (O'Neil, 2016). Scholars increasingly emphasize the importance of transparency, accountability, and fairness in fostering public trust in AI systems (Floridi et al., 2018; Binns, 2018; Mittelstadt et al., 2016).

While media interest and academic scrutiny of AI have grown, public perception remains highly heterogeneous. Moreover, the diversity of methodological approaches across studies complicates comparison and synthesis. In the following section, we examine risk and benefit perception as a conceptual framework for assessing public evaluations of emerging technologies, including AI.

2.2. Risk and benefit perception and the psychometric model

Human decision-making does not follow purely rational or utility-maximizing principles (Gigerenzer and Brighton, 2009). Rather than relying solely on objective calculations or classical economic models of decision-making (e.g., Neumann and Morgenstern, 1944; Bartholomae and Wiens, 2024), individuals' attitudes and judgments about risks and benefits are shaped by affective, cognitive, cultural, and emotional factors (Kahneman and Tversky, 1984; Witte and Allen, 2000; Hoffmann et al., 2015; Huang et al., 2020). This tendency is particularly evident in the context of emerging technologies such as artificial intelligence (AI), where public attitudes are often informed more by narratives, heuristics, and sociocultural framing than by technical understanding or direct experience (Cave et al., 2019). Consequently, public perceptions of the risks and benefits of AI often reflect broader individual or social hopes or anxieties, rather than an (impartial) assessment of empirical evidence. These perceptions, in turn, influence how individuals engage with, accept, or resist AI in everyday contexts.

Although risk can be formally defined in probabilistic terms—as the expected utility of adverse events and their potential consequences—this definition is of limited practical use in the context of emerging technologies and lay evaluations (Aven and Renn, 2009; Kahneman and Tversky, 1984; Fischhoff et al., 1978). Instead, researchers have emphasized the role of perceived risk, measured through subjective ratings of threat and concern. Slovic's *psychometric model* of risk perception (Slovic et al., 1979, 1986), along with related work (Fischhoff et al., 1978), provides a widely adopted framework that accounts for cognitive, affective, and experiential dimensions of risk judgment. This model has been applied across various domains, including genetic engineering (Connor and Siegrist, 2010), genetically modified food (Verdurme and Viaene, 2003), nuclear energy (Slovic et al., 2000), climate change (Pidgeon and Fischhoff, 2011), vaccination hesitance (Wong and Yang, 2021), and carbon capture technologies (Arning et al., 2020).

However, risk perception constitutes only one side of the evaluative process. Perceived benefit—defined as individuals' subjective assessment of positive outcomes associated with a technology—serves as the conceptual counterweight. Across many domains, research consistently demonstrates an strong but not perfect inverse relationship between perceived risks and perceived benefits (Alhakami and Slovic, 1994; Efendić et al., 2021): technologies viewed as risky tend to be perceived as offering fewer benefits, while those considered safe are typically seen as more beneficial. This inverse association underscores the importance of assessing both dimensions together to understand the psychological balancing process involved in technology evaluation. This dual-framework approach is particularly valuable for assessing emerging technologies whose implications are not yet fully understood and may be difficult for laypersons to evaluate using objective criteria. In such cases, early-stage public perceptions—however fragmented or incomplete—can offer valuable insights into potential societal responses and policy needs.

Recent empirical work supports the relevance of this framework in the AI context. In a study of AI applications in healthcare, both risk and benefit perceptions emerged as strong predictors of preferences for AI adoption (Kerstan et al., 2024). Trust in AI was found to reduce perceived risk, while objective knowledge was positively associated with perceived benefits. Similarly, Said et al. (2023) found that increased knowledge of AI may lead to a reduced perception of its risks, although this relationship is nuanced and potentially moderated by factors such as trust or media exposure. These findings align with the broader literature, which confirms the general inverse relationship between perceived risks and benefits for AI technologies (Alessandro et al., 2024). Importantly, this study also highlighted the role of application context: perceptions of AI varied significantly across scenarios involving law enforcement, propaganda, and entertainment, suggesting that domain-specific framings shape public evaluations of AI technologies.

2.3. AI in different contexts

Luhmann's *system theory* ("Systemtheorie") suggests six major subsystems (or functions) in today's society: economy, law, science, politics, religion, and education (Luhmann, 1989). Each subsystem fulfills a main function and cannot be replaced by another system (Peterson et al., 2004). Some researchers describe technology as a social system itself, but even without this approach, technology undoubtedly plays a huge role in today's society affecting each subsystem (Reichel, 2011). This view is reflected for AI research in the work of Liu (2021), in which the author states that research on AI can be categorized into three perspectives: scientific, technical and cultural AI research. The cultural perspective encompasses research on AI and its interaction with economic, cultural, social and political contexts, demonstrating that AI plays a role in multiple subsystems as proposed by Luhmann. This interaction and its characteristic of being a conglomerate of both older (digitalization, neural networks, machine learning, ...) and currently evolving technologies (big data, deep learning, large language models (Aghion et al., 2018; Surden, 2019; Van Noorden and Perkel, 2023; Tahiru, 2021) warrants a cross-sectional exploration of its impact.

The perception of AI is being explored and discussed in numerous individual contexts, scenarios, and therefore subsystems (Shinners et al., 2022; Henestrosa et al., 2023; Ragot et al., 2020), however, the current literature lacks a thorough understanding of how the perception of AI compares across the contexts and the scenario in which AI is envisioned.

Alessandro et al. (2024) conducted an extensive cross-context study on AI perception, covering 25 AI applications. The results showed that the perception of the applications varied significantly, from medical applications being perceived as highly risky with moderate value to society, to an intelligent chess game with low risk and value attributed, and political propaganda chatbots being perceived as highly valuable to society with minimal risk. Novozhilova et al. (2024) explored perceptions of ability and benevolence of AI across different domains such as healthcare, education, and creative arts, with healthcare often seen as more beneficial than the other domains.

With an experimental approach Liehner et al. (2021) studied the willingness to delegate morally sensitive tasks to automated AI-agents and found that, again, context and reliability (i.e., risk of an error) of the automation shapes the perception of and trust in AI. In this instance, reliance on the agents for the health context was lower compared to an urban and warehouse context.

Araujo et al. (2020) explored how individual differences relate to the perceptions of automated decision-making by AI, and how perceptions of risk, fairness, and usefulness of AI differ across three contexts: media, (public) health, and justice. While no difference was found for associated risk across the three contexts, the usefulness and fairness of automated decision-making by AI in the health and judicial context was perceived higher than in the media context. Furthermore, individual knowledge of AI had a positive impact on perceived usefulness and fairness of AI, while privacy concerns were associated with perceived risks.

Although these studies show that the perception of AI is indeed context-dependent, results are scarce and, especially in the context of health applications, contradictory. Therefore, we highlight the need to (a) further expand the set of contexts in which AI is assessed and compared, and (b) increase the number of studies making such a comparison to identify common themes. In addition to the influence of context, Araujo et al. (2020) showed that individual differences also play a role and are therefore explored in the next section.

2.3.1. Individual differences and sociocultural factors

A growing body of research suggests that individual differences—ranging from demographic characteristics such as age and gender to levels of AI-related knowledge and experience—significantly influence public perceptions of artificial intelligence (Yigitcanlar et al., 2022). In general, individuals with higher technological competence and greater familiarity with AI tend to express more trust in and acceptance of AI systems (Novozhilova et al., 2024; Crockett et al., 2020).

Personality traits have also been shown to shape AI attitudes. In a study conducted with a Turkish sample, Kaya et al. (2022) found that frequent computer use and greater knowledge about AI were associated with more positive attitudes towards the technology. Conversely, personality factors such as higher Agreeableness, AI learning anxiety, and configuration anxiety were linked to more negative attitudes.

Recent work by Winter et al. (2024) further underscores the importance of personality in AI adoption. Investigating ChatGPT usage, they found that perceived effectiveness and concerns about the technology predicted actual use frequency. Additionally, intention to use the tool was positively associated with Machiavellianism, a personality trait characterized by manipulative behavior and strategic thinking. These findings highlight that personality dimensions can influence not only general attitudes but also specific usage behaviors in AI contexts.

Cultural and national context is another factor in shaping AI perception. A large-scale cross-national survey of over 10,000 participants by Kelley et al. (2021) revealed notable geographic differences. Respondents from high-income countries such as the United States, Canada, and Australia often expressed ambivalence, combining concerns about AI's ethical implications with futuristic expectations. In contrast, participants from lower-income countries such as India, Brazil, and Nigeria tended to report greater enthusiasm and optimism regarding AI's potential. South Korean respondents emphasized AI's utility and practical applications, reflecting the country's high level of technological development. Despite regional variation, there was broad consensus that AI will have a profound societal impact, though the specific nature of this impact remains unclear.

These cultural patterns may be partly attributable to underlying social values. For instance, Kanzola et al. (2024) demonstrated that in a Greek sample, social identity factors such as altruism and openness to cultural change significantly influenced attitudes towards AI. Such findings support the broader claim that public perceptions of emerging technologies are mediated by both individual-level dispositions and broader sociocultural orientations.

2.4. *Technology forecasting through scenario-based methods and socio-technical imaginaries*

To better anticipate public responses to emerging technologies, researchers have increasingly turned to foresight methodologies and scenario analysis as tools for capturing uncertainty and exploring a range of plausible futures. Scenario-based approaches allow individuals to assess technologies not in abstract terms, but within concrete, contextualized use cases; thus enabling more nuanced reflections on potential societal impacts. Börjeson et al. (2006) propose a widely used typology distinguishing between predictive, explorative, and normative scenarios, each suited to different decision-making contexts. Popper (2008) further highlights the methodological diversity in foresight practices and emphasizes the value of participatory techniques in understanding stakeholder perspectives. As Veenman (2013) notes, scenario techniques are particularly useful in policymaking for engaging with complex and uncertain technological developments such as AI. Against this backdrop, scenario-based designs provide a structured framework for eliciting lay perceptions of AI, making them particularly relevant for early-stage assessments of societal readiness and governance needs.

2.4.1. *Research gap and questions*

In summary, while there is a growing body of literature examining public perceptions and use of AI, several critical gaps remain. One key limitation concerns the contextual specificity of AI perception: public attitudes often vary substantially depending on the domain of application, yet this variation remains insufficiently understood. Without a nuanced understanding of how people perceive AI across different contexts, policymakers and researchers face challenges in prioritizing areas for development, regulation, and public communication. Furthermore, understanding the underlying reasons for these perceptual differences is essential for tailoring AI technologies to user needs—an important factor in promoting acceptance and responsible adoption.

Against this backdrop, the present study investigates the following research questions:

1. In which domains and application areas is AI perceived as value-aligned (i.e., viewed positively or negatively)?
2. To what extent is overall sentiment towards AI driven by perceived benefits versus perceived risks?
3. Is the trade-off between perceived risks and benefits consistent across contexts, or does it vary depending on the application?
4. How do individual-level factors—such as age, gender, and technology-related personality traits—influence perceptions of AI in terms of risk, benefit, and overall value?

3. Method

This study examines public perceptions of AI, focusing on perceived benefits, risks, and overall value judgments. It also analyzes how individuals navigate trade-offs among these perceptions and investigates the influence of personal characteristics on these evaluations.

3.1. *Risk–benefit tradeoff using micro scenarios*

A common approach to studying technology perception involves asking participants to evaluate a specific scenario (or a limited set of scenarios) using comprehensive battery of rating scales (Veenman, 2013; Börjeson et al., 2006). While effective for analyzing specific applications, this method offers limited insight into general attitudes towards AI, given its wide-ranging applications and their individual, organizational, and societal implications. To address this, we adopted a multi-scenario approach in which participants evaluated a broad set of potential AI developments and societal impacts over the next decade, using brief, single-sentence micro scenarios (Brauner, 2024). Participants responded to short AI-related statements (e.g., "AI raises living standards") using a concise set of single-item scales.

This approach enables analysis from two distinct but complementary perspectives:

1. User-level analysis (individual differences): Each participant's average evaluations across all topics are treated as reflective indicators of an underlying latent construct. This allows examination of how individual differences—such as demographics or attitudes towards technology—influence overall AI perception.
2. Topic-level analysis (scenario-level analysis): By averaging responses across all participants for each scenario, we derive topic-specific perception scores. These can be visualized on risk–benefit maps and used to identify perception clusters, outliers, or areas of consensus.

To create the list of topics and statements, we drew on existing research literature and insights from workshops with domain experts. Through multiple rounds of refinement, we optimized the selection, eliminated redundancies, and revised the statements for clarity and conciseness. The topic selection also reflects elements of Luhmann's *systems theory* ("Systemtheorie"), which identifies six core subsystems—or functions—within modern society: economy, law, science, politics, religion, and education (Luhmann, 1989). The final set comprised 71 statements, ranging from widely accepted to more speculative claims—for example, that AI creates jobs, fosters innovation, operates according to moral principles, or perceives humans as a threat. Each participant evaluated a randomly selected subset of 15 scenarios, presented in randomized order to mitigate order effects and reduce cognitive fatigue. The full list of items is provided in Table A.1 in Appendix.

For the evaluation, we build on Slovic's psychometric model (Slovic et al., 1986) (see also Related Work). Rather than relying on objective cost-benefit analyses or probabilistic risk assessments, this study emphasizes participants' *subjective* perceptions of risk, benefit, and overall value. This approach is particularly well-suited to the present study for several reasons. First, AI, as an emergent technology, is characterized by uncertainty, complexity, and ethical ambiguity. Slovic's model is designed to capture how individuals intuitively process such factors through affective attributions. This is especially relevant given the diverse AI contexts examined, with different applications (e.g., autonomous weapons, predictive policing, autonomous driving) likely to evoke varying emotional responses. Second, the psychometric approach recognizes the interplay between perceived risks and perceived benefits, acknowledging that these dimensions, while related, are not strictly inverse (Alessandro et al., 2024). Third, the model enables the analysis of variation in risk and benefit perceptions across population groups. Given that AI perceptions differ by age, gender, and levels of technological readiness, a framework that supports systematic comparison across user characteristics enables a more nuanced understanding. Finally, Slovic's model has been validated across a wide range of technologies and societal risks, demonstrating its robustness, generalizability, and influence on real-world attitudes and behaviors (Wong and Yang, 2021; Shin et al., 2022).

Each topic was evaluated using five dependent variables measured on a single 6-point semantic differential item: (1) expectation of occurrence within the next decade (*will not happen—will happen*), (2) perceived personal risk (*low risk—high risk*), (3) perceived benefit (*useless—useful*), (4) perceived social risk (*socially harmful—socially harmless*), but excluded from the analysis (see below), and (5) general valuation or sentiment (*negative—positive*).[1]

The final item builds on the Value-based Adoption Model by Kim et al. (2007) and serves as the target variable for investigating how perceived risks and benefits shape overall evaluations of AI technology. The use of single-item measures is justified in this context, as the constructs are clearly defined and theoretically grounded (Rammstedt and Beierlein, 2014; Fuchs and Diamantopoulos, 2009), particularly drawing on prior work on the psychometric risk–benefit model (Slovic et al., 1979; Alhakami and Slovic, 1994).

3.2. Demographics and exploratory personality traits

In addition to the micro scenario evaluations, we collected demographic data from participants, including age (in years), gender (following Spiel et al. (2019) with the closed-choice options: male, female, diverse, and no response), current occupation, and highest level of educational attainment.

We also assessed several exploratory personality traits. Although we did not formulate specific hypotheses regarding the magnitude or direction of these effects, we expected these variables to influence both the perception of and attitudes towards AI.

Interpersonal Trust: Given that people tend to perceive technology as social actors (Fogg and Tseng, 1999; Reeves and Nass, 1996), we assume this perception may extend to AI, leading individuals to view it as a social actor. Since trust plays a critical role in mediating social relationships, we hypothesize that interpersonal trust is related to AI perception, and we measured it using the three-item KUSIV3 short scale (Nießen et al., 2021).

Technology Readiness (or Technology Commitment): This trait refers to an individual's propensity to embrace and effectively use new technologies. We hypothesize that higher technology readiness positively influences attitudes towards AI, and we assessed it using a subset of the Technology Commitment Scale (Neyer et al., 2016).

Openness: In the Big Five personality model, the trait openness is characterized by imagination, curiosity, and a preference for novelty, creativity, and diverse experiences. This trait has been linked to greater curiosity and more favorable attitudes towards technology, including AI (Kaya et al., 2022).

General Self-Efficacy: Self-efficacy refers to a person's belief in their ability to successfully perform tasks and handle challenges across various situations. Individuals with higher self-efficacy may feel more capable of understanding and engaging with AI technologies, which can positively shape their attitudes. We measured general self-efficacy using the General Self-Efficacy Short Scale-3 (GSE3) (Doll et al., 2021).

Risk Propensity: Risk propensity reflects an individual's tendency to take or avoid risks and indicates their comfort with uncertain outcomes. We expect that higher risk propensity is associated with more favorable attitudes towards AI, particularly regarding risk perception: individuals more tolerant of risk may view AI as a source of opportunity and innovation, while more risk-averse individuals may focus on potential threats such as job loss, privacy concerns, or loss of control. We measured risk propensity using a single-item scale (Nießen et al., 2020).

AI Readiness: Lastly, we assessed AI readiness using a subset of the Medical Artificial Intelligence Readiness Scale for Medical Students (MAIRS-MS; referred to here as AIRS) (Ozan Karaca and Demir, 2021). Although originally developed for medical contexts, we assume this scale is applicable in broader settings and may serve as a stronger predictor of positive attitudes towards AI than general technology readiness. We included it based on the assumption that greater AI-related knowledge and self-efficacy are associated with more favorable AI perceptions.

The survey opened with an informed consent form, informing participants that participation was voluntary, no personal data would be collected, and that the data would be made publicly available as open data. The questionnaire was administered in German. Fig. 1 illustrates the structure of the questionnaire.

3.3. Sample acquisition, data cleaning, and data analysis

The sample was recruited via an independent online research participant pool and was representative of the German population in terms of age, gender, and location. The study was approved by our university's institutional review board (IRB) under ID `2023_02b_FB7_RWTH Aachen`.

We analyzed the data using both parametric and non-parametric procedures, including the Bravais–Pearson correlation coefficient (r), Spearman's ρ, Chi-square (χ^2), Kendall's Tau (τ), and multiple linear regression analyses. We assessed the assumptions underlying each test and reported any violations. Missing responses were excluded on a per-test basis. In line with common practice in the social sciences, we set the Type I error rate at 5% (α = .05) for statistical significance (Andy Field, 2009).

We filtered the data to exclude incomplete or low-quality responses using the following criteria: the participant must (1) have fully completed the survey, (2) have passed the attention item (i.e., "please select 'rather agree'"), and (3) not have been classified as speeders (i.e., completing the survey in less than one-third of the median survey duration). These thresholds are typically sufficient for identifying low-quality responses in surveys (Leiner, 2019)). The median survey duration was 9.8 min. and the cutoff criterion therefore set at < 3.3 minutes. After filtering, the data included 1100 of original 1354 cases (dropout rate: 18.8%).

The scales demonstrated acceptable to high reliability: technology readiness (α = .883), interpersonal trust (α = .850), general self-efficacy (α = .830), openness from the Big Five model (α = .730), and the AI readiness scale AIRS (α = .920). We excluded the assessment dimension "perceived harmfulness" and focused on the dimension of perceived risk, as both dimensions are too highly correlated for meaningful inferences (r = .928, p < .001).

All materials, raw and unfiltered data, reproducible analyses, and assumption checks are publicly accessible in the open data repository on OSF (https://osf.io/gt9un/).

[1] In the Werturteilsstreit, Weber (1904) argued that science should remain objective and value-neutral, while acknowledging that values, norms, and ideals could themselves be valid subjects of research.

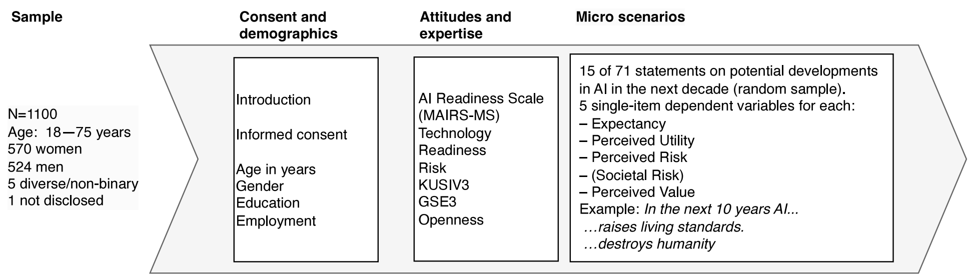

Fig. 1. Survey design: After obtaining informed consent, questions on demographics and explanatory user factors participants evaluated 15 out of 71 micro-scenarios related to potential AI capabilities.

3.4. Description of the sample

By using an online research participant pool, we ensured that the sample represents the population of Germany across key demographic variables such as age, gender, education, employment status, and geographical background. The final sample consists of 1100 participants (570 (51.8%) women; 524 (47.8%) men, 5 (0.5%) diverse or non-binary, 1 (0.1%) person did not disclose their gender identity). The age ranged from 18 to 85 years with a median age of 51 years. There is no association between age and gender in the sample ($\tau = -.031$, $p = .210 > .05$).

The participants in the sample report to have a diverse range of educational backgrounds. The majority of participants have completed their education at the university level, with 27.1% having an academic degree and 20.4% having a university entrance certificate (“Abitur” or “Fachabitur”). Another significant portion of participants completed a high school diploma (“Realschulabschluss”) (23.5%) or have vocational training (18.4%). A smaller percentage of participants have completed a secondary school certificate (“Hauptschulabschluss”, 10.5%), while only a few participants have no formal education (0.2%).

Participants reported different current employment statuses. The largest proportion of participants are currently employed full-time (48.2%), followed by those who are retired (22.5%). 14.8% of participants are employed part-time, while a smaller percentage are currently unemployed (7.8%) or in other employment relations, such as vocational training (0.7%), study programs (2.5%), or parental leave (1.3%). A very small percentage of participants are engaged in voluntary military or social services (0.1%), have irregular or mini jobs (1.9%), or are currently in school (0.2%). Overall, the sample consists of individuals with a wide range of employment statuses, reflecting different stages in their professional lives and personal circumstances. The characteristics of the sample are presented in Table 1.

In the sample, higher age was significantly associated with lower technical readiness ($r = -.224$, $p < .001$) and lower AI-readiness ($r = -.250$, $p < .001$), but not significantly with interpersonal trust ($r = .067$, $p = .114$), general self-efficacy ($r = .061$, $p = .131$), or openness ($r = -.071$, $p = .114$). Gender was associated with lower technology readiness scores ($r = -0.210$, $p < .001$) and AI-readiness ($r = -.156$, $p < .001$), with women reporting lower technology readiness and experience with AI, and openness ($r = .093$, $p = .014$), with women, on average, reporting slightly higher openness scores than men. No significant associations were found between gender and interpersonal trust (KUSIV3) ($r = .008$, $p = .779$), or gender and general self-efficacy ($r = -.048$, $p = .218$).

Table 1
Description of the nationally representative sample from Germany (N = 1100).

Variable	N	Percent
Age		
18–85 years (median: 51 years)	1100	100.0%
Gender		
Female	570	51.8%
Male	524	47.6%
Diverse	5	0.5%
Not Specified	1	0.1%
Education		
No formal education	2	0.2%
Lower secondary school diploma	115	10.5%
Intermediate secondary school diploma	259	23,5%
Higher education entrance qualification	224	20,4%
University degree	298	27,1%
Vocational Training	202	18.4%
Employment		
Full time employed	530	48,2%
Part time employed	163	14,8%
Early retiree, retiree, pensioner	248	22.5%
Currently unemployed	86	7.8%
Student	27	2,5%
other	46	4.2%

4. Results

First, we present average evaluations across the four assessment dimensions—*Expectancy*, *Perceived Risk*, *Perceived Benefit*, and overall attributed *Value*—aggregated across topics and participants. These averages offer insight into the general public perception of AI. Second, we analyze how individual topics were rated across these four dimensions and examine how the dimensions relate to one another. Finally, we investigate individual-level trade-offs between perceived risks and benefits, with a focus on how user diversity and personality-related differences shape AI evaluations.

4.1. Overall assessment of AI

On average, participants rated the expectancy that these AI developments will materialize as slightly above neutral (+12.7%), suggesting a general belief that most projections are likely to become reality within the next decade. Perceived risk is relatively high (+34.7%), as shown by the gray distribution in Fig. 2; only a few topics are viewed as relatively safe, falling below the neutral point. Perceived benefit is close to neutral (−5.2%), but with considerable variation: some AI projections are seen as highly beneficial, while others are considered largely unhelpful. Overall attributed value (or general sentiment towards AI) is somewhat negative (−19.7%), though a subset of topics received clearly positive evaluations. Fig. 2 illustrates these distributions, and the left panel of Table 2 summarizes the average scores across all topics and participants.

4.2. Evaluations of the queried AI statements

Given the large number of topics included in the study, we do not provide a detailed discussion of each individual statement and its evaluation here. Readers interested in the full set of results can refer to Table A.1 in the Appendix, which lists the average ratings for all items.

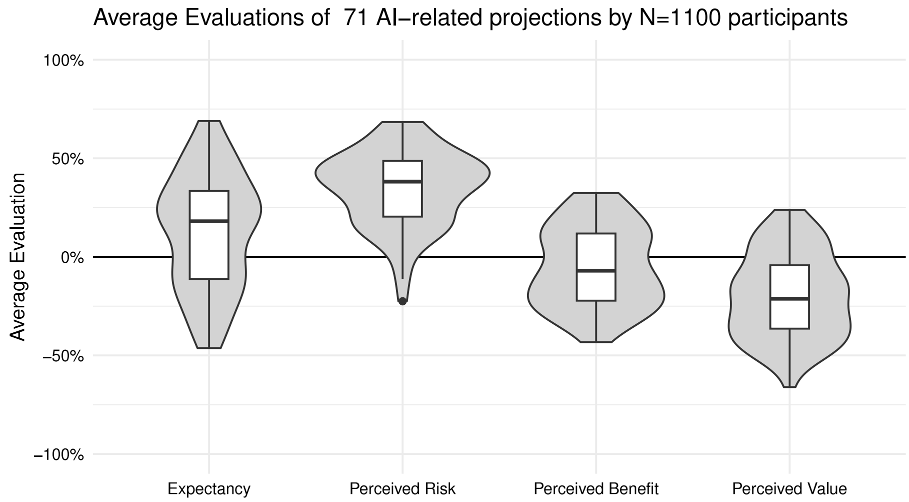

Fig. 2. Average evaluation of the 71 micro scenarios on the four assessment dimensions *Expectancy*, *Perceived Risk*, *Perceived Benefit*, and overall *Value*. N = 1100 participants as box- and violin plot. The gray area illustrates the distribution of the topic evaluations regarding the respective dimension.

Instead, we focus on reporting the three highest- and lowest-rated topics for each of the four evaluation dimensions.

Regarding the anticipated likelihood of AI-related developments, the three most expected statements were: "AI will independently drive automobiles" (+68.9%), "AI is misused by criminals" (+67.7%), and "AI learns faster than humans" (+60.5%). Conversely, the three least expected statements were: "AI helps us to have better relationships" (−46.3%), "AI is a family member" (−43.0%), and "AI has a sense of responsibility" (−40.7%).

For perceived risk, the most concerning statements were: "AI is misused by criminals" (+68.4%), "AI supervises our private life" (+67.1%), and "AI determines warfare" (+66.4%). On the other end of the spectrum, the safest-rated statements were: "AI is humorous" (−22.5%), "AI creates valuable works of art that are traded for money" (−11.2%), and "AI serves as a conversation partner in elderly care" (−5.9%).

Concerning the perceived benefit of the various AI-related concepts, the three most positively rated statements were: "AI carries out medical diagnoses" (+32.3%), "AI promotes innovation" (+31.7%), and "AI improves our health" (+31.1%). In contrast, the statements considered least useful were: "AI reduces our need for interpersonal relationships" (−36.7%), "AI creates valuable works of art that are traded for money" (−37.6%), and "AI decides about our death" (−43.2%).

Finally, participants rated the overall value (or valence or sentiment) of each projection—i.e., whether they attributed a more positive or negative overall value to it (Kim et al., 2007). The most positively valued statements were: "AI improves our health" (+23.8%), "AI serves as a conversation partner in elderly care" (+18.1%), and "AI supports me as a helper in my tasks" (+17.6%). The most negatively rated statements were: "AI is misused by criminals" (−66.1%), "AI decides about our death" (−51.4%), and "AI supervises our private life" (−50.6%).

While this univariate analysis offers a foundational understanding of participants' perceptions of AI, we now shift to a bivariate perspective, examining the relationships between the evaluation dimensions through correlation analyses, followed by regression modeling.

4.2.1. Relationships among the topic evaluations

We examined whether evaluations of the AI-related topics across the different assessment dimensions were interrelated. The right panel of Table 2 displays the correlations among these dimensions.

The average expectancy that a given AI-related development will occur within the next decade was not significantly associated with its perceived risk (r = .150, p = .632), perceived benefit (r = .277, p = .077), or overall value (r = .054, p = .449). This suggests that the anticipated likelihood of AI projections becoming reality neither drives nor is substantially influenced by participants' risk, benefit, or value evaluations. In other words, the perceived temporal distance or feasibility of AI developments does not appear to affect how risky or beneficial they are seen to be.

Table 2
Correlation matrix showing strong associations among perceived risk, perceived benefit, and overall value (sentiment) of the N = 71 topic evaluations. Expectancy, i.e., whether participants believe the depicted AI scenarios are likely to occur within the next decade, shows no significant correlations with the other dimensions. Asterisks (***) indicate statistically significant correlations at p < .001; values in parentheses represent non-significant correlations.

	M	SD	Expectancy	Risk	Benefit	Value
Expectancy	12.7%	66.7%	–	(+0.150)	(+0.277)	(+0.054)
Perceived Risk	34.7%	56.4%		–	−0.524 ***	−0.800 ***
Perceived Benefit	−5.2%	59.0%			–	+0.904***
Overall Value	−19.7%	57.8%				–

In contrast, perceived risk was negatively correlated with both perceived benefit ($r = -.524$, $p < .001$) and overall value ($r = -.800$, $p <$ .001). That is, topics perceived as more risky were also viewed as less beneficial and received more negative overall evaluations. Conversely, perceived benefit was strongly positively correlated with overall value (r = +.904, p < .001), indicating that AI applications considered more useful were also evaluated more positively (and vice versa).

Since both perceived risk (negatively) and perceived benefit (positively) influence the overall attributed value of the topics—and given that risk and benefit themselves are interrelated—we conducted a multiple linear regression analysis. This approach allowed us to disentangle and quantify the unique contributions of perceived risk and perceived benefit in predicting the overall perceived value of each topic, which served as the dependent variable.

The regression model, which included both perceived risk (β = −.490, $p < .001$) and perceived benefit ($\beta = +.672$, $p < .001$), identified both predictors as strong and statistically significant contributors to the overall perceived value. In contrast, the interaction term between risk and benefit ($\beta = +.138$, $p = .305$), as well as the intercept ($I = 0.014$, $p =$ .265), were not statistically significant. The model exhibited excellent fit ($R^2 = .965$, $F(3, 67) = 612.3$, $p < .001$), indicating that perceived risk and benefit together explain the vast majority of the variance in overall sentiment. Importantly, multicollinearity was not a concern, with all *VIF* values below 1.5. Table 3 presents the detailed regression coefficients. These results highlight that, even when accounting for their intercorrelation, both risk and benefit independently and substantially shape the perceived value of AI.

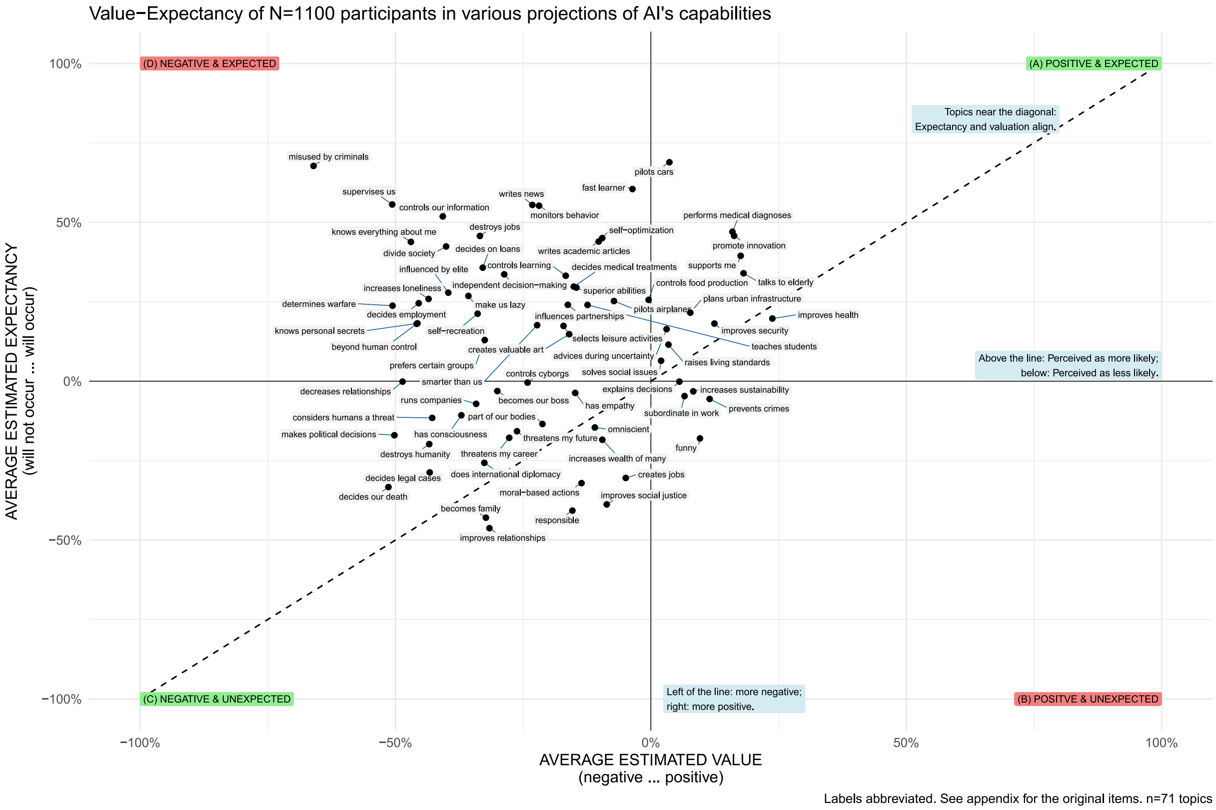

Fig. 3. Average evaluation of the 71 micro scenarios along the dimensions of *Perceived Value* (x-axis) and *Expectancy* (y-axis), based on responses from $N = 1100$ participants. While there is no significant correlation between the two assessment dimensions across topics ($r = .054$, $p > .999$), a substantial number of scenarios fall into the quadrant representing outcomes that are expected yet evaluated negatively.

Table 3
Multiple regression analysis predicting the average overall value of the 71 AI-related topics from perceived risk and perceived benefit. The model is statistically significant and explains 96.5% of the variance in overall value ($R^2 = .965$).

Variable	Std. β = B	SE	T	p
(Intercept)	0.012	0.011	1.125	.265
Perceived Risk	−0.490	0.033	−14.722	< .001
Perceived Benefit	0.672	0.046	14.564	< .001
Perc. Risk × Benefit	0.138	0.134	1.035	.305

4.2.2. *Expectancy and perceived value of the queried topics*

Fig. 3 visualizes all 71 AI-related projections, plotting their average perceived value on the x-axis against their expectancy, i.e., how likely participants believe the projection is to materialize within the next decade, on the y-axis. The chart can be interpreted as follows: Horizontally, points on the left represent AI projections evaluated more negatively, while points on the right indicate more positive assessments. Vertically, points near the top reflect projections perceived as more likely to occur, whereas those closer to the bottom are seen as less likely. This results in four meaningful interpretive quadrants: (a) positive and expected developments, (b) positive but unexpected projections, (c) negative and unlikely scenarios, and (d) negative yet expected outcomes.

Additionally, points that lie along or near the diagonal suggest an alignment between expectancy and evaluation: participants perceive these projections as both likely and positive or unlikely and negative, indicating a degree of consensus. In contrast, points that deviate substantially from this diagonal reveal areas of tension or controversy (and suggest areas with research potential and governance needs): projections that are considered likely but evaluated negatively, or unlikely yet viewed positively. The overall scattered distribution underscores the absence of a systematic relationship between expectancy and perceived value, reinforcing our earlier finding that the anticipated likelihood of AI developments does not predict their overall evaluation.

4.2.3. *Risk–benefit tradeoff of the queried topics*

Fig. 4 maps each of the 71 AI-related projections according to participants' assessments of perceived risk (x-axis) and perceived benefit (y-axis). As with previous figure, this visualization can be interpreted via four quadrants: Horizontally, projections positioned towards the left are seen as less risky, while those on the right are perceived as more risky. Vertically, projections placed lower on the axis are considered to offer limited benefit, whereas those higher up are seen as more useful. This results in four interpretive zones: (a) high risk/high benefit—projections that may be transformative yet controversial, (b) high risk/low benefit—technologies largely viewed as both dangerous and of limited benefit, (c) low risk/low benefit, benign but unremarkable applications, and (d) low risk/high benefit, safe and beneficial innovations.

As the figure shows, relatively few AI-related projections are perceived as low-risk, reflected by the sparsity of points on the left side of the diagram. For example, the idea that AI can create valuable art is seen as largely harmless but also of limited practical benefit, placing it in the quadrant of benign yet marginally useful technologies. Similarly, the notion that AI can be humorous is perceived as safe but of neutral benefits. In contrast, the projection that AI could act as a conversational partner in elderly care stands out as both low-risk and highly beneficial, positioning it as a strong candidate for positive public adoption.

Most other AI projections are distributed along a continuum from moderate to high perceived risk, with varying levels of attributed benefit. Unlike the previous figure (depicting expectancy vs. sentiment), Fig. 4 reveals a pronounced negative relationship between risk and benefit, illustrated by the clustering of topics along the upward-sloping regression line; mirroring the earlier statistical findings. This suggests

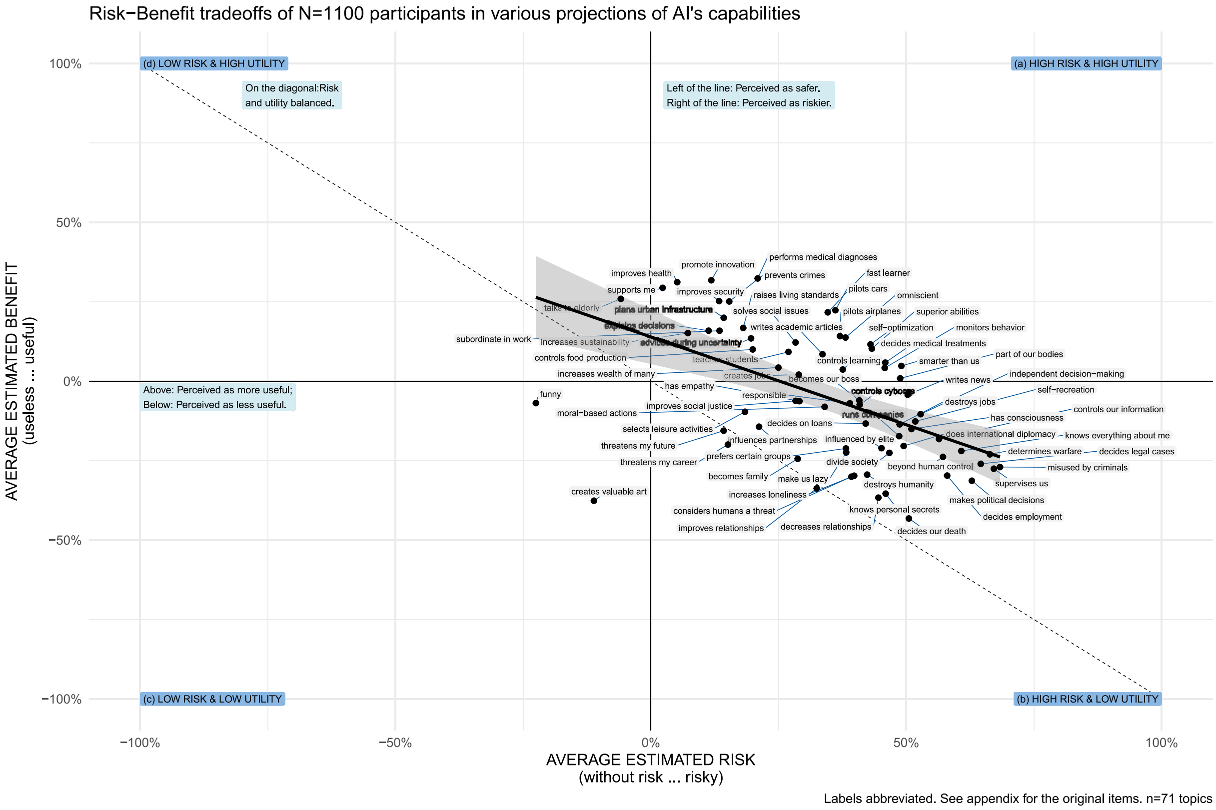

Fig. 4. Average evaluation of the 71 AI-related projections in terms of perceived risk (x-axis) and perceived benefit (y-axis) by $N = 1100$ participants. Across topics, risk and benefit ratings are strongly and negatively correlated ($r = -.524$, $p < .001$), indicating that projections viewed as more risky tend to be seen as less beneficial, and vice versa. The black line represents the linear regression fit, with the gray shaded area indicating the 95% confidence interval.

that, in public perception, potential utility and perceived threat tend to be inversely related, highlighting the complex trade-offs inherent in emerging AI technologies.

4.3. Perception of AI as individual difference

In the following, we interpret the perception of AI as an individual difference variable and investigate how demographic characteristics and personality traits influence participants' evaluations. To this end, we treat participants' average ratings of the AI topics across the four evaluative dimensions—*expectancy*, *perceived risk*, *perceived benefit*, and overall *value*—as reflective indicators of the corresponding latent constructs (Brauner, 2024). This approach enables us to assess how psychological and socio-demographic factors influence individual-level variation in the perception of AI.

As the lower part of Table 4 illustrates, several demographic and personality factors are significantly associated with participants' evaluations of the AI projections. Expectancy, i.e., the perceived likelihood that the AI projections will materialize within the next decade, is positively associated with participants' openness from the Big Five personality model ($r = .136$, $p < .001$) and their AI readiness (AIRS; $r = .094$, $p = .036$). Participants with higher openness judged the projections as more likely. A weak positive trend also exists for AI readiness, indicating that greater hands-on exposure may elevate expectations. Perceived risk is positively associated with age ($r = .197$, $p < .001$) and negatively associated with both technology readiness ($r = -.152$, $p < .001$) and AI readiness ($r = -.175$, $p < .001$). In other words, older individuals tend to perceive AI as more risky, whereas those who are more comfortable with or knowledgeable about technology and AI tend to see it as less threatening. Perceived benefit is likewise associated with age ($r = -.182$, $p < .001$), technology readiness ($r = .233$, $p < .001$), and AI readiness ($r = .274$, $p < .001$). Older participants tend to evaluate AI as less beneficial, while higher levels of readiness correlate with more favorable views. Interestingly, neither gender nor general self-efficacy was significantly related to any of the four evaluation dimensions.

Table 4
Significant correlations of demographic and attitudinal factors with the individual perceived expectancy, risk, benefit, and value towards the AI projections (N = 1100), "." signifies insignificant correlations).

Variable	AI expectancy	AI risk	AI benefit	AI value
AI Expectancy	–	+.212	+.143	.
AI Risk		–	−.639	−.711
AI Benefit			–	+.869
Age in Years	.	+.197	−.182	−.146
Gender (dummy coded m = 1, w = 2)	.	.	.	.
Openness (Big 5)	+.136	.	.	.
General Self-Efficacy (GSE3)	.	.	.	.
Technology Readiness	.	−.152	+.233	+.207
AI Readiness (AIRS/MAIRS-MS)	+.094	−.175	+.274	+.223

In addition, the upper part of Table 4 shows the coupled inter-relationships among the four evaluation dimensions (interpreted as personality factor): Expectancy is weakly but significantly associated with perceived risk ($r = .212$, $p < .001$) and benefit ($r = .143$), but not with overall value. This supports earlier findings from the topic-wise perspective that expectancy is decoupled from how AI topics are emotionally or morally evaluated. Again, perceived risk and perceived benefit are strongly negatively correlated ($r = -.639$, $p < .001$), indicating a clear tradeoff pattern. Finally, the individuals' overall value judgments are strongly influenced by both risk ($r = -.749$, $p < .001$) and benefit ($r = +.869$, $p < .001$).

To investigate how demographic and user-related factors influence perceived AI risk, perceived benefit, and the overall attributed value

Table 5
Hierarchical linear regression results predicting perceived risk, perceived benefit, and attributed value of AI based on demographics and technology attitudes. Including Technology Readiness and AI Readiness improved the models' explanatory power and decreased the influence of age (and to a lesser extent, gender) on the three target variables (N = 1.094). "***" significant at $p < .001$, "**" significant at $p < .01$, "*" significant at $p < .05$.

Independent variable	Perceived risk	Perceived benefit	Perceived value
Step 1: Demographics			
(Intercept)	+0.074	+0.214***	+0.059
Age in Years (β)	+0.198***	−0.183***	−0.149***
Gender (β, dummy coded m = 1, w = 2)	+0.051	−0.055	−0.068***
R^2	.041	.036	.026
$F(2, 1091)$	23.46***	20.19***	14.65***
Step 2: Explanatory Variables			
(Intercept)	+0.332***	+0.286***	+0.390***
Age in Years (β)	+0.159***	−0.109***	−0.088*
Gender (β, dummy coded m = 1, w = 2)	+0.022	+0.000	+0.021
Technology Readiness (β)	+0.058	+0.109*	+0.108*
AI Readiness (AIRS) (β)	−0.100**	+0.189***	+0.141***
ΔR^2	+.018	+.061	+.042
R^2	.059	.097	.067
$F(4, 1089)$	16.92***	29.22***	19.85***

of AI as dependent variables, we conducted hierarchical multiple regression analyses using two sequential blocks of predictors. In the first block, we included the demographic variables *age* and *gender* to assess baseline sociodemographic effects. In the second block, we added *technology readiness* and *AI readiness* as psychological user characteristics, based on their theoretical relevance and prior empirical associations with technology perception. Other potential predictors were excluded from the model due to their lack of significant bivariate correlations with the dependent variables (see Table 4).

In the first block of the hierarchical regression analyses, results indicated that all three models were statistically significant: perceived AI risk ($F(2, 1091) = 23.46$, $p < .001$, $R^2 = .041$), perceived AI benefit ($F(2, 1091) = 20.19$, $p < .001$, $R^2 = .036$), and overall attributed AI value ($F(2, 1091) = 14.65$, $p < .001$, $R^2 = .026$). Across all three models, *age* emerged as a consistent and significant predictor: older participants reported higher levels of perceived AI risk ($\beta = +.198$), lower perceived benefit ($\beta = -.183$), and a more negative overall evaluation of AI ($\beta = -.149$). In contrast, *gender* was only a significant predictor for overall AI value ($\beta = +.068$), indicating that women attributed slightly lower overall value to AI than men, while gender did not significantly affect perceived risk or benefit.

The second-level models, which included both technology readiness and AI readiness as predictors, were also statistically significant and demonstrated improved model fit across all three dependent variables: perceived AI risk ($F(4, 1089) = 16.92$, $p < .001$, $R^2 = .059$), perceived AI benefit ($F(4, 1089) = 29.22$, $p < .001$, $R^2 = .097$), and overall AI value ($F(4, 1089) = 19.85$, $p < .001$, $R^2 = .068$). The inclusion of AI and technology readiness in the second block significantly enhanced the explanatory power of each model. Specifically, higher levels of AI readiness were associated with lower perceived risk, greater perceived benefit, and a more positive overall evaluation of AI. Technology readiness, while not significantly related to perceived risk, emerged as a significant positive predictor of both perceived benefit and overall value. Importantly, with the inclusion of these psychological readiness factors, the effect of age diminished across all models, and the previously observed gender effect on overall AI value was rendered non-significant. Table 5 summarizes the hierarchical regression results in detail.

4.4. Desired foci of AI governance

Lastly, we asked participants to identify what they viewed as the primary focus for effective AI governance (single-choice response). According to their responses, the most important priority (selected by 45.3% of participants) was ensuring human control and supervision over AI development and use. Other key areas included transparency (13.0%), data protection and data management (11.7%), and promoting social and ecological well-being (9.3%). Participants also cited diversity, non-discrimination, and fairness (4.8%), robustness and security (4.7%), and accountability (4.5%) as relevant, though less dominant, concerns. In total, 6.7% of respondents did not answer this question.

5. Discussion

Artificial Intelligence (AI) may become one of the defining technologies of the 21st century. Understanding how people perceive and balance its risks and benefits is crucial for ensuring that AI research and implementation align with human values and support effective governance. Drawing on the psychometric paradigm (Slovic et al., 1986), we explored how individuals assess trade-offs between perceived risks and benefits across various AI-related micro-scenarios, considering both individual and technological perspectives. Our analysis is based on a representative sample of 1100 participants from Germany.

Across the various topics surveyed, we observed predominantly negative sentiment: Most topics were perceived as relatively risky for individuals and of limited personal benefit. This pattern may partly reflect a bias in topic selection, as many statements highlighted negative or challenging scenarios (such as "AI will be misused by criminals" or "AI determines warfare"). However, even seemingly positive or neutral statements, such as "AI creates many jobs" or "AI will independently drive automobiles", elicited caution. Participants tended to attribute higher risks to these scenarios, regardless of their potential perceived benefits. Overall, fewer than 20% of the statements received a positive evaluation, and many of those were still considered risky.

Beyond absolute evaluations, we also analyzed how overall sentiment is formed. First, we observed an inverse relationship between perceived risks and perceived benefits. This finding supports previous research on AI perception (Alessandro et al., 2024) as well as broader studies on risk perception across various contexts (Alhakami and Slovic, 1994; Efendić et al., 2021). While this relationship may appear intuitive, it reflects a cognitive bias: individuals tend to downplay the benefits of a technology they perceive as risky (or vice versa). This bias can shape public attitudes and influence policy decisions regarding technology adoption.

We also found that overall sentiment (ranging from negative to positive) is shaped by both perceived risks and perceived benefits, with benefits exerting a stronger influence. On one hand, higher perceived risk leads to more negative emotional responses to AI. On the other hand, perceived benefits have a significantly positive impact, often outweighing the negative effects of risk in shaping overall attitudes. This aligns with earlier findings in risk research, which suggest that perceptions of technology are primarily driven by perceived benefits rather than risks (Alhakami and Slovic, 1994; Efendić et al., 2021). This is particularly notable given that AI is often viewed negatively or with concern, especially in many Western countries (Cave et al., 2019; Curran et al., 2019; Kelley et al., 2021). Overall, our results suggest that clearly demonstrating the benefits of AI is essential to fostering more positive public attitudes. However, effective risk management remains critical, as high perceived risks can undermine even strong perceived benefits.

From the perspective of individual differences, our results suggest that both demographic factors and individual attitudes shape perceptions of AI's risks, benefits, and overall evaluation. Younger respondents tended to view AI-related topics as less risky, more beneficial, and assigned them a higher overall value. Gender also played a role—though to a lesser extent—with women generally giving lower evaluations of AI than men. These findings are consistent with current research on AI perception, which indicates that age and gender influence attitudes towards AI (e.g., Yigitcanlar et al. (2022), Crockett et al. (2020)). Moreover, individuals with higher levels of technological or AI literacy tend to be less apprehensive about AI (Novozhilova et al., 2024).

However, these demographic effects diminish among individuals with higher levels of technology or AI readiness. This suggests that increasing technology and AI literacy could help reduce perceived risks, increase perceived benefits, and improve overall acceptance of AI in society. When people better understand the basic functioning of AI, the ethical challenges involved in building sophisticated models, and the broader implications of AI for individuals, organizations, and society, it becomes possible to engage in a robust, democratic debate about AI's potential, limitations, and consequences at various societal levels. Such informed dialog can provide a foundation for deciding where AI should assume control, where it should serve in a supportive role, and which areas ought to remain AI-free. Efforts like free online courses for adults, such as "Elements of AI", and updated school curricula that incorporate digitalization and AI literacy are essential (Olari and Romeike, 2021; Marx et al., 2022).

Still, even after accounting for technology and AI readiness, a significant age-related bias remained. Given that AI development is largely driven by younger, mostly male developers, this raises concerns about whether current and future AI systems will reflect the values and norms of the broader population of users (Young et al., 2023; Brauner et al., 2024). To address this, it is vital to educate developers in human-centered and participatory design methods and to integrate ethical training into computer science education and AI development processes.

The visual maps offer a common ground for identifying critical topics, discussing research needs, and exploring the societal implications of AI. These cartographies of AI perceptions help charting areas considered more critical than others and represent a foundation that future research can build upon.

Beyond enhancing our understanding of public perceptions of AI, this article also contributes methodologically. By using micro-scenarios (Brauner, 2024), we examined public perceptions from two distinct perspectives: as expressions of individual differences and as evaluations of technology. This dual approach is valuable because it provides a more nuanced understanding of how people perceive AI, not only capturing general attitudes but also uncovering factors that vary across individuals. These insights are particularly relevant for policymakers and technology developers. They demonstrate that public opinion on AI is multifaceted, reflecting a balance between perceived risks and benefits, often with benefits outweighing concerns about risk. Understanding these dynamics can support the development of more targeted public communication strategies tailored to different usage contexts and audiences, as well as regulatory approaches that resonate with public concerns while emphasizing AI's practical value. This methodological approach thus enriches theoretical perspectives and offers practical guidance for aligning AI development and governance with public sentiment.

5.1. Implications for policy and societal futures

These findings offer actionable insights for AI governance and societal foresight. Since perceived benefits have a stronger influence than perceived risks in shaping public evaluations, policy communication should emphasize concrete, relatable examples of how AI can enhance everyday life. At the same time, domains associated with high perceived risk, such as AI in warfare, political decision-making, or surveillance, require targeted regulation, public deliberation, and ethical oversight.

Public perception will also influence societal adoption. If AI is seen as threatening or lacking usefulness, its integration into healthcare, public services, or the workplace may encounter significant resistance. Therefore, forward-looking governance must foster trust through participatory design processes, transparency requirements, and domain-specific regulations. The risk–benefit cartographies introduced in this study can help policymakers identify areas of concern and prioritize interventions. As AI continues to shape lives, workplaces, and societies, success and alignment will depend not only on technical feasibility, but also on public legitimacy and perceived value.

5.2. Limitations

This study is not without limitations. First, although the participant sample was large and diverse in terms of age, gender, education, and employment, we surveyed only individuals from Germany. Given existing research that suggests cultural differences in AI perception (Kelley et al., 2021; Curran et al., 2019), future studies should extend this work by analyzing how cultural dimensions, such as country of origin or cultural heritage, influence AI perceptions and the associated risk–benefit trade-offs.

Second, although the selection of AI-related topics was informed by prior research, it may still be subject to bias, potentially resulting in spurious findings due to Berkson's paradox (Berkson, 1946). Future work could address this by systematically designing a broader topic set or focusing on underexplored areas. Despite this limitation, the alignment between topic evaluations (perspective 2) and individual sentiment towards AI (perspective 1) suggests a consistent pattern. Because we used a representative sample, any topic selection bias is unlikely to translate into bias in individual differences. This supports the generalizability of the findings and suggests the topic selection was sufficiently robust.

Third, and most importantly, the study examined many statements about AI's future implications using a limited number of dependent variables. This approach captures participants' heuristic evaluations rather than in-depth, cognitively demanding assessments. Recent studies show that survey results can be influenced not only by respondents' judgments, but also by linguistic properties of the items themselves, such as word co-occurrence (Gefen and Larsen, 2017). To address this, our study employed reflexive measurements across a diverse range of topics rather than relying solely on similarly phrased psychometric scales (Brauner, 2024). Despite the heuristic nature of responses, our results reveal strong and systematic evaluation patterns, indicating that the method yields reliable insights. This suggests the approach offers a novel perspective for triangulating cognitive phenomena in technology perception. Still, while our findings provide a broad overview of public sentiment and its link to individual differences, they offer limited insight into the specific motivations behind individual evaluations of particular topics. As AI continues to shape individual lives and society, each topic warrants deeper qualitative and quantitative investigation to inform researchers, practitioners, and policymakers aiming to better align AI with human needs.

Fourth, recent anthropological research highlights how human identity is dynamically shaped through ongoing interactions with technology (Hasse, 2022). The widespread integration of AI into everyday life has the potential to challenge and redefine self-perception and societal norms. As such, our study captures a contemporary snapshot shaped by current cultural contexts and norms. Future research should explicitly consider these evolving human-technology dynamics to better understand the existential dimensions of public attitudes towards AI—and how these perceptions may vary across cultures and shift over time.

6. Conclusion and outlook

This study investigated how people in Germany perceive the risks, benefits, and societal value of artificial intelligence across a wide range of domains. Based on public evaluations of 71 AI-related scenarios, we found that while many AI applications are viewed as likely to occur, they are often associated with high perceived risk and limited benefit. Crucially, regression analyses revealed that perceived benefits play a stronger role than perceived risks in shaping how people value AI, while perceived likelihood has little to no influence.

These findings have practical implications for system design, public communication, and AI governance. To foster societal acceptance and ethical alignment, AI developers and policymakers should prioritize enhancing the perceived benefits of AI—particularly in domains currently viewed with skepticism—while also addressing public concerns and fears. Moreover, individual differences such as age, gender, and especially AI readiness significantly shaped participants' perceptions. This highlights the importance of targeted education, public AI literacy, and inclusive, participatory approaches in shaping equitable AI futures.

In light of these findings, we echo calls by scholars such as Shneiderman (2021) and Crawford (2021) to move beyond purely technical narratives and towards human-centered, value-aligned AI. AI is not neutral: it is embedded in social structures, resource use, and labor systems, reflecting human assumptions and choices. Our study contributes to this discourse by offering empirically grounded maps of public perception, which can inform more responsible and democratically informed AI development.

From a regulatory perspective, our findings highlight important gaps in current frameworks. The EU's 2024 "Artificial Intelligence Act" (AI Act) provides a robust structure to mitigate harmful AI practices by classifying applications into risk categories and imposing proportionate requirements (European Union, 2024). However, while the Act offers strong protections against unacceptable and high-risk AI (that are either banned or need to comply with security and transparency regulations), it lacks concrete mechanisms for promoting AI systems that align with public values and enhance well-being.

Given our finding that perceived benefits are the strongest predictor of AI value assessments, this gap in the AI Act is particularly consequential. Without clear incentives or guidelines for maximizing public benefit, even legally compliant AI systems may fail to gain acceptance and their positive potential may remain unrealized. We therefore call for complementary, value-focused design guidelines that go beyond compliance, encouraging developers to create systems that actively promote human well-being, trust, and social alignment.

Our data suggest that the public currently holds a risk-oriented view of AI: most projections were evaluated as riskier than safe, even those that would generally be considered socially beneficial, such as job creation, social cohesion, or decision support under uncertainty. This underscores the need not just to mitigate harm, but to better communicate the value and usefulness of AI in areas where it can make a positive impact.

One particularly concerning gap is the use of AI in warfare, which is explicitly excluded from the scope of the AI Act, yet was among the scenarios rated as most risky in our study. Although the EU recognizes the importance of human control and legal accountability in military AI applications (Sebastian, 2025), comprehensive legislation remains absent. Other scenarios, such as the effect of AI on relationships or personal identity, also fall outside current regulatory frameworks, despite their deep social relevance.

As Awad et al. (2018) point out, it is not enough for AI systems to comply with legal or ethical standards, they also require social legitimacy. Conversely, public preferences may not always align with what is ethically sound or legally permissible. In this light, our findings should not be read as prescriptive guidance for bans or specific regulations. Rather, they can help identify domains of public concern, inform research priorities, and inspire extra-regulatory frameworks that promote human-centered AI development. As Valdez et al. (2024) warns, without robust methods to evaluate AI's societal impact, "the EU AI Act may lead to repeating the mistakes of the GDPR and to rushed, chaotic, ad-hoc, and ambiguous implementation".

Looking ahead, we strongly encourage cross-cultural and longitudinal studies to monitor how public perceptions evolve over time and across different societies. We also call for more nuanced research comparing risk–benefit tradeoffs across domains and use cases. Ultimately, integrating public values early in the design and regulation of AI is not just a normative goal; it is a prerequisite for building systems that are both socially legitimate and broadly beneficial.

CRediT authorship contribution statement

Philipp Brauner: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Felix Glawe:** Writing – review & editing, Validation, Formal analysis. **Gian Luca Liehner:** Writing – review & editing, Writing – original draft, Validation, Methodology, Conceptualization. **Luisa Vervier:** Writing – review & editing, Validation, Methodology, Investigation, Conceptualization. **Martina Ziefle:** Writing – review & editing, Validation, Supervision, Software, Project administration, Funding acquisition.

Ethics approval

Our universities IRB approved this study under ID `2023_02b_FB7_RWTH Aachen`

Consent to participate

Participant were informed about the goal and approach of the study and that data will be stored on a public repository. They gave informed consent.

Consent for publication

All authors provided their consent for publication.

Code availability

All materials, data, and analysis are available on OSF: https://osf.io/gt9un/

Use of AI

During the preparation of this manuscript, the authors used ChatGPT for language editing and coding support during data analysis. All content was subsequently reviewed and revised by the authors, who take full responsibility for the integrity and accuracy of the final publication.

Funding

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC- 2023 Internet of Production – 390621612. Open Access funding enabled and organized by Projekt DEAL.

Declaration of competing interest

We affirm that this manuscript has not been published elsewhere and is not under consideration by any other publication. All authors have approved the submission and declared no conflicts of interest.

Acknowledgments

This study builds upon the conceptual foundation laid by a previous collaboration between Ralf Philipsen and the first author (Brauner et al., 2023). The first author gratefully acknowledges the shared development of the initial idea and the inspiring discussions that led to this work. The methodology employed here evolved through several studies and was shaped by numerous conversations with colleagues. We would like to especially thank Julia Offermann, Julian Hildebrandt, and Ralf Philipsen for their valuable input on the research design, data analysis, and interpretation of findings. Our sincere thanks also go to Tim Schmeckel for his dedicated research assistance throughout the project. We are grateful to Anonymous Reviewer 2 for the rapid review process and highly constructive feedback. Support of this quality and speed is rare and greatly appreciated.

Appendix

Table Table A.1 lists the average expectancy, risk, benefit, and value evaluations for all queried topics, ordered by value.

Data availability

Data is freely available (materials, data, analyses): https://osf.io/6q3eu/.

Table A.1
Question items from the survey and average responses to each item from the participants (N = 1100) ordered by attributed value.

In 10 years, AI will...	Expectancy	Risk	Benefit	Value
be misused by criminals.	67.7%	68.4%	−27.0%	−66.1%
decide about our death.	−33.3%	50.5%	−43.2%	−51.4%
supervise our private life.	55.6%	67.1%	−27.6%	−50.6%
determine warfare.	23.8%	66.4%	−23.0%	−50.6%
make political decisions.	−17.0%	62.8%	−31.3%	−50.2%
reduce our need for interpersonal relationships.	−0.1%	44.5%	−36.7%	−48.6%
know everything about me.	43.8%	60.8%	−22.0%	−47.0%
no longer be controllable by humans.	18.1%	57.2%	−23.8%	−45.8%
know my secrets.	18.2%	46.0%	−35.4%	−45.7%
decide on hiring, promotions, and terminations.	24.5%	58.0%	−29.7%	−45.5%
lead to personal loneliness.	25.9%	39.8%	−29.8%	−43.5%
destroy humanity.	−19.8%	42.3%	−29.4%	−43.4%
administer justice in legal matters.	−28.7%	64.6%	−26.0%	−43.3%
consider humans as a threat.	−11.6%	39.3%	−30.1%	−42.8%
control what messages we receive.	51.9%	56.4%	−18.2%	−40.7%
divides society.	42.4%	46.7%	−22.6%	−40.1%
be influenced by an elite.	27.8%	45.1%	−21.1%	−39.7%
have its own consciousness.	−10.7%	51.0%	−15.1%	−37.1%
make society lazy.	26.8%	38.2%	−22.4%	−35.7%
run companies.	−7.1%	48.6%	−17.3%	−34.2%
recreate itself.	21.2%	51.8%	−12.7%	−33.9%
destroy many jobs.	45.7%	48.7%	−13.6%	−33.5%
decide who gets an important financial loan.	35.7%	42.1%	−13.4%	−32.9%
conduct international diplomacy.	−25.7%	49.5%	−20.4%	−32.6%
prefer certain groups of people.	13.0%	38.2%	−21.2%	−32.5%
be a family member.	−43.0%	28.7%	−24.4%	−32.3%
help us to have better relationships.	−46.3%	32.5%	−33.8%	−31.6%
occupy leadership positions in working life.	−3.1%	40.8%	−6.1%	−30.1%
make independent decisions that affect our lives.	33.6%	52.8%	−10.4%	−28.7%
threaten my professional future.	−17.8%	15.1%	−20.0%	−27.7%
threaten my private future.	−15.8%	14.2%	−15.6%	−26.2%
control hybrids of humans and technology.	−0.4%	40.9%	−7.4%	−24.1%
independently write news.	55.5%	50.3%	−4.3%	−23.2%
be more intelligent than humans.	17.6%	49.0%	4.8%	−22.3%
supervise our behavior in public.	55.2%	45.8%	4.1%	−21.9%
become part of the human body.	−13.5%	48.8%	0.9%	−21.2%
determine our leisure time activities.	17.4%	18.4%	−9.6%	−17.1%
control what and how we learn.	33.2%	37.6%	3.7%	−16.7%
control our search for partners.	24.0%	21.2%	−14.3%	−16.3%
create valuable works of art that are traded for money.	14.8%	−11.2%	−37.6%	−16.0%
have a sense of responsibility.	−40.7%	39.0%	−7.0%	−15.4%
decide on medical treatments.	29.7%	45.9%	5.8%	−15.1%
have the ability to recognize, understand and empathize with emotions.	−3.7%	29.1%	−6.3%	−14.8%
be ahead of humans in its abilities.	29.5%	43.3%	10.2%	−14.6%
act according to moral concepts.	−32.1%	34.0%	−8.1%	−13.6%
teache students.	24.0%	26.9%	9.2%	−12.4%
be omniscient.	−14.6%	38.1%	13.7%	−11.0%
independently write scientific articles.	44.0%	33.6%	8.5%	−10.2%

(continued on next page)

Table A.1 (*continued*).

In 10 years, AI will...	Expectancy	Risk	Benefit	Value
be able to optimize itself.	45.0%	43.0%	11.6%	−9.5%
increase the wealth of many people.	−18.4%	25.0%	4.3%	−9.5%
increase social justice.	−38.8%	28.3%	−6.2%	−8.6%
autonomously take off, fly, and land airplanes.	25.2%	37.0%	14.2%	−7.2%
create many jobs.	−30.5%	29.0%	2.0%	−4.9%
learn faster than humans.	60.5%	34.6%	21.6%	−3.6%
control food production.	25.6%	19.9%	10.0%	−0.4%
contribute to solving complex social problems.	6.4%	28.3%	12.2%	2.0%
advise me in uncertain times.	16.4%	19.6%	13.4%	3.1%
raise our standard of living.	11.5%	18.1%	16.7%	3.4%
independently drive automobiles.	68.9%	36.1%	22.3%	3.6%
explain its decisions.	−0.1%	11.3%	15.9%	5.6%
always be subordinate to us in working life.	−4.7%	13.4%	15.9%	6.6%
determine the construction and infrastructure of our cities.	21.5%	14.3%	19.9%	7.7%
make our society more sustainable.	−3.2%	7.2%	15.1%	8.3%
be humorous.	−18.0%	−22.5%	−6.9%	9.6%
prevent crimes.	−5.6%	15.3%	25.1%	11.5%
improve the security of people.	18.1%	13.4%	25.2%	12.4%
carry out medical diagnoses.	47.0%	20.9%	32.3%	16.0%
promotes innovation.	45.7%	11.8%	31.7%	16.3%
support me as a helper in my tasks.	39.4%	2.3%	29.4%	17.6%
serve as a conversation partner in elderly care.	34.0%	−5.9%	25.9%	18.1%
improve our health.	19.7%	5.2%	31.1%	23.8%
AVERAGE	**12.7%**	**34.7%**	**−5.2%**	**−19.7%**

Note: Measured on 6 point semantic differentials and rescaled to −100% to +100%. Negative values indicate a negative evaluation of the dimension (i.e., low value, low perceived risk, low perceived benefit, or low expectancy) and positive values indicate a high evaluation. Permission to translate, use, and adapt the items is—of course—granted.

References

Acemoglu, Daron, Restrepo, Pascual, 2017. The race between machine and man. Am. Econ. Rev. 108 (6), 1488–1542. http://dx.doi.org/10.3386/w22252.

Aghion, Philippe, Jones, Benjamin F., Jones, Charles I., 2018. Artificial Intelligence and Economic Growth. in the Economics of Artificial Intelligence: An Agenda. University of Chicago Press, pp. 237–282.

Alessandro, Gabbiadini, Federica, Durante, Cristina, Baldissarri, Luca, Andrighetto, 2024. Artificial intelligence in the eyes of society: Assessing social risk and social value perception in a novel classification. Hum. Behav. Emerg. Technol. (ISSN: 2578-1863) 2024, 1–11. http://dx.doi.org/10.1155/2024/7008056.

Alhakami, Ali Siddiq, Slovic, Paul, 1994. A psychological study of the inverse relationship between perceived risk and perceived benefit. Risk Anal. (ISSN: 1539-6924) 14 (6), 1085–1096. http://dx.doi.org/10.1111/j.1539-6924.1994.tb00080.x.

Amunts, Katrin, Axer, Markus, Banerjee, Swati, Bitsch, Lise, Bjaalie, Jan G., Brauner, Philipp, Brovelli, Andrea, Calarco, Navona, Carrere, Marcel, Caspers, Svenja, Charvet, Christine J., Cichon, Sven, Cools, Roshan, Costantini, Irene, D'Angelo, Egidio Ugo, De Bonis, Giulia, Deco, Gustavo, DeFelipe, Javier, Destexhe, Alain, Dickscheid, Timo, Diesmann, Markus, Düzel, Emrah, Eickhoff, Simon B., Einevoll, Gaute, Eke, Damian, Engel, Andreas K., Evans, Alan C., Evers, Kathinka, Fedorchenko, Nataliia, Forkel, Stephanie J., Fousek, Jan, Friederici, Angela D., Friston, Karl, Furber, Stephen, Geris, Liesbet, Goebel, Rainer, Güntürkün, Onur, Abd Hamid, Aini Ismafairus, Herold, Christina, Hilgetag, Claus C., Hölter, Sabine M., Ioannidis, Yannis, Jirsa, Viktor, Kashyap, Sriranga, Kasper, Burkhard S., D'Exaerde, Alban de Kerchove, Kooijmans, Roxana, Koren, István, Kotaleski, Jeanette Hellgren, Kiar, Gregory, Klijn, Wouter, Klüver, Lars, Knoll, Alois C., Krsnik, Zeljka, Kämpfer, Julia, Larkum, Matthew E., Linne, Marja-Leena, Lippert, Thomas, Abdullah, Jafri Malin, Di Maio, Paola, Magielse, Neville, Maquet, Pierre, Mascaro, Anna Letizia Allegra, Marinazzo, Daniele, Mejias, Jorge, Meyer-Lindenberg, Andreas, Migliore, Michele, Michael, Judith, Morel, Yannick, Morin, Fabrice O., Muckli, Lars, Nagels, Guy, Oden, Lena, Palomero-Gallagher, Nicola, Panagiotaropoulos, Fanis, Paolucci, Pier Stanislao, Pennartz, Cyriel, Peeters, Liesbet M., Petkoski, Spase, Petkov, Nicolai, Petro, Lucy S., Petrovici, Mihai A., Pezzulo, Giovanni, Roelfsema, Pieter, Ris, Laurence, Ritter, Petra, Rockland, Kathleen, Rotter, Stefan, Rowald, Andreas, Ruland, Sabine, Ryvlin, Philippe, Salles, Arleen, Sanchez-Vives, Maria V., Schemmel, Johannes., Senn, Walter, de Sousa, Alexandra A., Ströckens, Felix, Thirion, Bertrand, Uluda, Kâmil, Vanni, Simo, van Albada, Sacha Jennifer, Vanduffel, Wim, Vezoli, Julien, Vincenz-Donnelly, Lisa, Walter, Florian, Zaborszky, Laszlo, 2023. Imaging Neuro Sci. 2024, 1–35. http://dx.doi.org/10.1162/imag_a_00137.

Andy Field, 2009. Discovering Statistics using SPSS, third ed. Sage Publications, ISBN: 1847879071.

Anon, 2023. The alan turing institute and ada Lovelace institute. Underst. Public Attitudes To Ai URL https://www.turing.ac.uk/. (Accessed 06 November 2024).

Araujo, Theo, Helberger, Natali, Kruikemeier, Sanne, de Vreese, Claes H., 2020. In ai we trust? perceptions about automated decision-making by artificial intelligence.. AI & SOCIETY (ISSN: 1435-5655) 35 (3), 611–623. http://dx.doi.org/10.1007/s00146-019-00931-w.

Arning, Katrin, Heek, Julia Offermann-van, Sternberg, André, Bardow, André, Ziefle, Martina, 2020. Risk-benefit perceptions and public acceptance of carbon capture and utilization. Environ. Innov. Soc. Transitions (ISSN: 2210-4224) 35, 292–308. http://dx.doi.org/10.1016/j.eist.2019.05.003.

Aven, Terje, Renn, Ortwin, 2009. On risk defined as an event where the outcome is uncertain. J. Risk Res. (ISSN: 1466-4461) 12 (1), 1–11. http://dx.doi.org/10.1080/13669870802488883.

Awad, Edmond, Dsouza, Sohan, Kim, Richard, Schulz, Jonathan, Henrich, Joseph, Shariff, Azim, Bonnefon, Jean-François, Rahwan, Iyad, 2018. The moral machine experiment. Nature (ISSN: 0028-0836) 1. http://dx.doi.org/10.1038/s41586-018-0637-6.

Bartholomae, Florian, Wiens, Marcus, 2024. Decision Theory. Springer Fachmedien Wiesbaden, ISBN: 9783658446086, pp. 3–31. http://dx.doi.org/10.1007/978-3-658-44608-6_1.

Berkson, Joseph, 1946. Limitations of the application of fourfold table analysis to hospital data, vol. 2, pp. 47–53. http://dx.doi.org/10.2307/3002000.

Binns, Reuben, 2018. Fairness in machine learning: Lessons from political philosophy. In: Friedler, Sorelle A., Wilson, Christo (Eds.), In: Proceedings of the 1st Conference on Fairness, Accountability and Transparency, vol. 81, PMLR, pp. 149–159.

Börjeson, Lena, Höjer, Mattias, Dreborg, Karl-Henrik, Ekvall, Tomas, Finnveden, Göran, 2006. Scenario types and techniques: Towards a user's guide. Futures (ISSN: 0016-3287) 38 (7), 723–739. http://dx.doi.org/10.1016/j.futures.2005.12.002.

Bostrom, Nick (Ed.), 2003. Superintelligence: Paths, Dangers, Strategies. Oxford University Press.

Bouschery, Sebastian G., Blazevic, Vera, Piller, Frank T., 2023. Augmenting human innovation teams with artificial intelligence: Exploring transformer-based language models. J. Prod. Innov. Manage. (ISSN: 1540-5885) 40 (2), 139–153. http://dx.doi.org/10.1111/jpim.12656.

Brauner, Philipp, 2024. Mapping acceptance: Micro scenarios as a dual-perspective approach for assessing public opinion and individual differences in technology perception. Front. Psychol. (ISSN: 1664-1078) 15, http://dx.doi.org/10.3389/fpsyg.2024.1419564.

Brauner, Philipp, Dalibor, Manuela, Jarke, Matthias, Kunze, Ike, Koren, István, Lakemeyer, Gerhard, Liebenberg, Martin, Michael, Judith, Pennekamp, Jan, Quix, Christoph, Rumpe, Bernhard, van der Aalst, Wil, Wehrle, Klaus, Wortmann, Andreas, Ziefle, Martina, 2022. A computer science perspective on digital transformation in production. ACM Trans. Internet Things (ISSN: 2691-1914) 3 (2), 1–32. http://dx.doi.org/10.1145/3502265.

Brauner, Philipp, Glawe, Felix, Liehner, Gian Luca, Vervier, Luisa, Ziefle, Martina, 2024. (Mis)alignments in AI perception: Quantitative findings and visual mapping of how experts and the public differ in expectations and risks, benefits, and value judgments. URL https://arxiv.org/abs/2412.01459.

Brauner, Philipp, Hick, Alexander, Philipsen, Ralf, Ziefle, Martina, 2023. What does the public think about artificial intelligence?—A criticality map to understand bias in the public perception of AI. Front. Comput. Sci. (ISSN: 2624-9898) 5, http://dx.doi.org/10.3389/fcomp.2023.1113903.

Brynjolfsson, Erik, McAfee, Andrew, 2014. The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. New York.

Budish, Ryan, 2021. Ai's risky business: Embracing ambiguity in managing the risks of ai. J. Bus. Technol. Law 16, 259.
Cath, Corinne, 2018. Governing artificial intelligence: ethical, legal and technical opportunities and challenges. Philosophical transactions of the royal society a:. Math. Phys. Eng. Sci. (ISSN: 1471-2962) 376 (2133), 20180080. http://dx.doi.org/10.1098/rsta.2018.0080.
Cave, Stephen, Coughlan, Kate, Dihal, Kanta, 2019. Scary robots: Examining public responses to ai. In: Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society. AIES '19, New York, http://dx.doi.org/10.1145/3306618.3314232, ACM.
Chen, Lijia, Chen, Pingping, Lin, Zhijian, 2020. Artificial intelligence in education: A review. IEEE Access 8, 75264–75278. http://dx.doi.org/10.1109/ACCESS.2020.2988510.
Connor, Melanie, Siegrist, Michael, 2010. Factors influencing people's acceptance of gene technology: The role of knowledge, health expectations, naturalness, and social trust. Sci. Commun. (ISSN: 1552-8545) 32 (4), 514–538. http://dx.doi.org/10.1177/1075547009358919.
Crawford, Kate, 2021. Atlas of ai: power, politics, and the planetary costs of artificial intelligence. In: Yale University Press.
Crockett, Keeley, Garratt, Matt, Latham, Annabel, Colyer, Edwin, Goltz, Sean, 2020. Risk and trust perceptions of the public of artifical intelligence applications. In: 2020 International Joint Conference on Neural Networks. IJCNN, IEEE, pp. 1–8. http://dx.doi.org/10.1109/ijcnn48605.2020.9207654.
Curran, Nathaniel Ming, Sun, Jingyi, Hong, Joo-Wha, 2019. Anthropomorphizing alphago: a content analysis of the framing of google deepmind's alphago in the chinese and american press. AI & SOCIETY 35 (3), 727–735. http://dx.doi.org/10.1007/s00146-019-00908-9.
Deng, Jia, Dong, Wei, Socher, Richard, Li, Li-Jia, Li, Kai, Fei-Fei, Li, 2009. Imagenet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. pp. 248–255. http://dx.doi.org/10.1109/CVPR.2009.5206848.
Diakopoulos, Nicholas, 2019. Automating the News: How Algorithms are Rewriting the Media. Harvard University Press, ISBN: 9780674976986.
Doll, E.S., Nießen, Désirée, Schmidt, I., Rammstedt, B., Lechner, C.M., 2021. General self-efficacy short scale-3 (gse-3). In: ZIS - the Collection of Items and Scales for the Social Sciences. http://dx.doi.org/10.6102/ZIS294, URL https://zis.gesis.org/DoiId/zis294.
Efendić, Emir, Chandrashekar, Subramanya Prasad, Lee, Cheong Shing, Yeung, Lok Yan, Kim, Min Ji, Lee, Ching Yee, Feldman, Gilad, 2021. Risky therefore not beneficial: Replication and extension of finucane, others,ś (2000) affect heuristic experiment. Soc. Psychol. Pers. Sci. (ISSN: 1948-5514) 13 (7), 1173–1184. http://dx.doi.org/10.1177/19485506211056761.
European Union, 2024. Regulation (EU) 2024/1244 of the European parliament and of the council of 13 2024 laying down harmonised rules on artificial intelligence (artificial intelligence act). Off. J. Eur. Union, L 94, 162024, https://eur-lex.europa.eu/eli/reg/2024/1244/oj.
Fast, Ethan, Horvitz, Eric, 2017. Long-term trends in the public perception of artificial intelligence. In: AAAI'17: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. pp. 963–969.
Fischhoff, Baruch, Slovic, Paul, Lichtenstein, Sarah, Read, Stephen, Combs, Barbara, 1978. How safe is safe enough? a psychometric study of attitudes towards technological risks and benefits. Policy Sci. (ISSN: 1573-0891) 9 (2), 127–152. http://dx.doi.org/10.1007/bf00143739.
Floridi, Luciano, Cowls, Josh, Beltrametti, Monica, Chatila, Raja, Chazerand, Patrice, Dignum, Virginia, Luetge, Christoph, Madelin, Robert, Pagallo, Ugo, Rossi, Francesca, Schafer, Burkhard, Valcke, Peggy, Vayena, Effy, 2018. Ai4people—an ethical framework for a good ai society: Opportunities, risks, principles, and recommendations. Minds Mach. (ISSN: 1572-8641) 28 (4), 689–707. http://dx.doi.org/10.1007/s11023-018-9482-5.
Fogg, B.J., Tseng, Hsiang, 1999. The elements of computer credibility. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '99. ACM, New York, NY, USA, pp. 80–87. http://dx.doi.org/10.1145/302979.303001.
Forsythe, Diana E., 1993. Engineering knowledge: The construction of knowledge in artificial intelligence. Soc. Stud. Sci. 23 (3), 445–477.
Friedman, Batya, Nissenbaum, Helen, 1996. Bias in computer systems. ACM Trans. Inf. Syst. (ISSN: 1046-8188) 14 (3), 330–347. http://dx.doi.org/10.1145/230538.230561.
Fuchs, Christoph, Diamantopoulos, Adamantios, 2009. Using single-item measures for construct measurement in management research – conceptual issues and application guidelines. Die Betriebswirtschaft 9.
Gabriel, Iason, 2020. Artificial intelligence, values, and alignment. Minds Mach. (ISSN: 1572-8641) 30 (3), 411–437. http://dx.doi.org/10.1007/s11023-020-09539-2.
Garcia, Max, 2024. Keeping the Box Black: AI Regulation and Corporate Ambiguity. (Ph.D. thesis). University of Oregon.
Gefen, David, Larsen, Kai, 2017. Controlling for lexical closeness in survey research: A demonstration on the technology acceptance model. J. Assoc. Inf. Syst. (ISSN: 1536-9323) 18 (10), 727–757. http://dx.doi.org/10.17705/1jais.00469.
Gigerenzer, Gerd, Brighton, Henry, 2009. Homo heuristicus: Why biased minds make better inferences. Top. Cogn. Sci. (ISSN: 17568757) 1 (1), 107–143. http://dx.doi.org/10.1111/j.1756-8765.2008.01006.x.
Gursoy, Furkan, Kakadiaris, Ioannis A., 2023. Artificial intelligence research strategy of the united states: critical assessment and policy recommendations. Front. Big Data (ISSN: 2624-909X) 6, http://dx.doi.org/10.3389/fdata.2023.1206139.
Hasse, Cathrine, 2022. Humanism, Posthumanism, and New Humanism: How Robots Challenge the Anthropological Object: Posthumanism. Springer Nature Singapore, ISBN: 9789811670848, pp. 145–164. http://dx.doi.org/10.1007/978-981-16-7084-8_7.
Henestrosa, Angelica Lermann, Greving, Hannah, Kimmerle, Joachim, 2023. Automated journalism: The effects of ai authorship and evaluative information on the perception of a science journalism article. Comput. Hum. Behav. (ISSN: 0747-5632) 138, 107445. http://dx.doi.org/10.1016/j.chb.2022.107445.
Hoffmann, Arvid O.I., Post, Thomas, M.E., Joost, 2015. Pennings. How investor perceptions drive actual trading and risk-taking behavior. J. Behav. Financ. (ISSN: 1542-7579) 16 (1), 94–103. http://dx.doi.org/10.1080/15427560.2015.1000332.
Holzinger, Andreas, Fister, Iztok, Kaul, Hans-Peter, Asseng, Senthold, 2024. Human-centered ai in smart farming: toward agriculture 5.0. IEEE Access (ISSN: 2169-3536) 12, 62199–62214. http://dx.doi.org/10.1109/access.2024.3395532.
Hopfield, John J., 1982. Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. (ISSN: 1091-6490) 79 (8), 2554–2558. http://dx.doi.org/10.1073/pnas.79.8.2554.
Hristova, Tsvetelina, Magee, Liam, Soldatic, Karen, 2024. The problem of alignment. AI & SOCIETY (ISSN: 1435-5655) http://dx.doi.org/10.1007/s00146-024-02039-2.
Huang, Xingyu, Dai, Shanshan, Xu, Honggang, 2020. Predicting tourists' health risk preventative behaviour and travelling satisfaction in tibet: Combining the theory of planned behaviour and health belief model. Tour. Manag. Perspect. (ISSN: 2211-9736) 33, 100589. http://dx.doi.org/10.1016/j.tmp.2019.100589.
Ipsos, 2022. Global attitudes towards artificial intelligence. Ipsos Report.
Jungherr, Andreas, 2023. Artificial intelligence and democracy: A conceptual framework. Soc. Media Soc. (ISSN: 2056-3051) 9 (3), http://dx.doi.org/10.1177/20563051231186353.
Kahneman, Daniel, Tversky, Amos, 1984. Choices, values, and frames. Am. Psychol. (ISSN: 0003-066X) 39 (4), 341–350. http://dx.doi.org/10.1037/0003-066x.39.4.341.
Kanzola, Anna-Maria, Papaioannou, Konstantina, Petrakis, Panagiotis, 2024. Unlocking society's standings in artificial intelligence. Technol. Forecast. Soc. Change (ISSN: 0040-1625) 200, 123106. http://dx.doi.org/10.1016/j.techfore.2023.123106.
Kaya, Feridun, Aydin, Fatih, Schepman, Astrid, Rodway, Paul, Yetişensoy, Okan, Kaya, Meva Demir, 2022. The roles of personality traits, ai anxiety, and demographic factors in attitudes toward artificial intelligence. Int. J. Human Comput. Interact. (ISSN: 1532-7590) 40 (2), 497–514. http://dx.doi.org/10.1080/10447318.2022.2151730.
Kelley, Patrick Gage, Yang, Yongwei, Heldreth, Courtney, Moessner, Christopher, Sedley, Aaron, Kramm, Andreas, Newman, David T., Woodruff, Allison, 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. AIES '21, ACM, http://dx.doi.org/10.1145/3461702.3462605.
Kerstan, Sophie, Bienefeld, Nadine, Grote, Gudela, 2024. Choosing human over ai doctors? how comparative trust associations and knowledge relate to risk and benefit perceptions of ai in healthcare. Risk Anal. 44 (4), 939–957. http://dx.doi.org/10.1111/risa.14216, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/risa.14216.
Kim, Hee-Woong, Chan, Hock Chuan, Gupta, Sumeet, 2007. Value-based adoption of mobile internet: an empirical investigation. Decis. Support Syst. (ISSN: 0167-9236) 43 (1), 111–126. http://dx.doi.org/10.1016/j.dss.2005.05.009.
Lecun, Yann, Bengio, Yoshua, Hinton, Geoffrey, 2015. Deep learning. Nature (ISSN: 14764687) 521 (7553), 436–444. http://dx.doi.org/10.1038/nature14539.
Leiner, Dominik Johannes, 2019. Too fast, too straight,too weird: Non-reactive indicators for meaningless data in internet surveys. Surv. Res. Methods 13 (3), 229–248. http://dx.doi.org/10.18148/srm/2019.v13i3.7403.
Liehner, Gian Luca, Brauner, Philipp, Schaar, Anne Kathrin, Ziefle, Martina, 2021. Delegation of moral tasks to automated agents the impact of risk and context on trusting a machine to perform a task. IEEE Trans. Technol. Soc. (ISSN: 2637-6415) 3 (1), 46–57. http://dx.doi.org/10.1109/TTS.2021.3118355.
Liu, Zheng, 2021. Sociological perspectives on artificial intelligence: A typological reading. Sociol. Compass (ISSN: 1751-9020) 15 (3), http://dx.doi.org/10.1111/soc4.12851.
Luhmann, Niklas, 1989. Ecological Communication. University of Chicago Press.
Makridakis, Spyros, 2017. The forthcoming artificial intelligence (ai) revolution: its impact on society and firms. Futures (ISSN: 0016-3287) 90, 46–60. http://dx.doi.org/10.1016/j.futures.2017.03.006.
Marx, Erik, Leonhardt, Thiemo, Baberowski, David, Bergner, Nadine, 2022. Using matchboxes to teach the basics of machine learning: an analysis of (possible) misconceptions. In: M., Katherine, Kinnaird, Peter Steinbach, Guhr, Oliver (Eds.), In: Proceedings of the Second Teaching Machine Learning and Artificial Intelligence Workshop, vol. 170, PMLR, pp. 25–29, URL https://proceedings.mlr.press/v170/marx22a.html.
McCarthy, John, Minsky, Marvin L., Rochester, Nathaniel, Shannon, Claude E., 2006. A proposal for the dartmouth summer research project on artificial intelligence. AI Mag. 27 (4), 12–14.

Mittelstadt, Brent Daniel, Allo, Patrick, Taddeo, Mariarosaria, Wachter, Sandra, Floridi, Luciano, 2016. The ethics of algorithms. Big Data Soc. (ISSN: 2053-9517) 3 (2), http://dx.doi.org/10.1177/2053951716679679.
Neumann, John Von, Morgenstern, Oskar, 1944. Theory of Games and Economic Behavior. University Press, Princeton, USA.
Neyer, Franz J., Felber, Juliane, Gebhardt, Claudia, 2016. Kurzskala technikbereitschaft (tb, technology commitment). Zusammenstellung Sozialwissenschaftlicher Items Und Skalen (ZIS) http://dx.doi.org/10.6102/ZIS244.
Nießen, Désirée, Beierlein, Constanze, Rammstedt, Beatrice, Lechner, Clemens M., 2021. Interpersonal trust short scale (kusiv3). ZIS – the collection of items and scales for the social sciences. http://dx.doi.org/10.6102/ZIS292_EXZ.
Nießen, Désirée, Groskurth, Katharina, Rammstedt, Beatrice, Lechner, Clemens M., 2020. Risk proneness short scale (r-1). ZIS Collection Items Scales Social Sci. http://dx.doi.org/10.6102/ZIS286.
Nissenbaum, Helen, 2001. How computer systems embody values. Computer (ISSN: 0018-9162) 34 (3), 118–120. http://dx.doi.org/10.1109/2.910905.
Novozhilova, Ekaterina, Mays, Kate, Paik, Sejin, Katz, James E., 2024. More capable, less benevolent: Trust perceptions of ai systems across societal contexts. Mach. Learn. Knowl. Extr. (ISSN: 2504-4990) 6 (1), 342–366. http://dx.doi.org/10.3390/make6010017.
Olari, Viktoriya, Romeike, Ralf, 2021. Addressing ai and data literacy in teacher education: A review of existing educational frameworks. In: The 16th Workshop in Primary and Secondary Computing Education, WiPSCE '21. ACM, http://dx.doi.org/10.1145/3481312.3481351.
O'Neil, Cathy, 2016. Weapons of math destruction. In: How Big Data Increases Inequality and Threatens Democracy. Crown Publishing Group, USA, ISBN: 0553418815.
Ozan Karaca, S. Ayhan Çalı şkan, Demir, Kadir, 2021. Medical artificial intelligence readiness scale for medical students (mairs-ms) – development, validity and reliability study. BMC Med. Educ. (ISSN: 1472-6920) 21 (1), http://dx.doi.org/10.1186/s12909-021-02546-6.
Peterson, Tarla, Peterson, Markus, Grant, William, 2004. Chapter two: social practice and biophysical process. In: The Environmental Communication Yearbook, vol. 1, pp. 15–32. http://dx.doi.org/10.1207/s15567362ecy0101_2.
Pew Research Center, 2023. What americans know about everyday uses of artificial intelligence. URL https://www.pewresearch.org/. (Accessed 06 November 2024).
Pidgeon, Nick, Fischhoff, Baruch, 2011. The role of social and decision sciences in communicating uncertain climate risks. Nat. Clim. Chang. (ISSN: 1758-6798) 1 (1), 35–41. http://dx.doi.org/10.1038/nclimate1080.
Popper, Rafael, 2008. The Handbook of Technology Foresight, chapter Foresight Methodology, firt ed. Edward Elgar, United Kingdom, ISBN: 978-1-84542-586-9, pp. 44–88.
Ragot, Martin, Martin, Nicolas, Cojean, Salomé, 2020. Ai-generated vs. human artworks. a perception bias towards artificial intelligence? In extended abstracts of the 2020 CHI conference on human factors in computing systems. pp. 1–10.
Rammstedt, Beatrice, Beierlein, Constanze, 2014. Can't we make it any shorter? the limits of personality assessment and ways to overcome them. J. Individ. Differ. (ISSN: 2151-2299) 35 (4), 212–220. http://dx.doi.org/10.1027/1614-0001/a000141.
Reeves, Byron, Nass, Clifford, 1996. The Media Equation – how People Treat Computers, Television, and New Media Like Real People and Places. Cambridge University Press.
Reichel, André, 2011. Technology as system: towards an autopoietic theory of technology. Int. J. Innov. Sustain. Dev. (ISSN: 1740-8822) 5 (2–3), 105–118. http://dx.doi.org/10.1504/IJISD.2011.043070.
Rumelhart, David E., Hinton, Geoffrey E., Williams, Ronald J., 1986. Learning representations by back-propagating errors. Nature (ISSN: 1476-4687) 323 (6088), 533–536. http://dx.doi.org/10.1038/323533a0.
Sadek, Malak, Calvo, Rafael A., Mougenot, Céline, 2024. Closing the socio–technical gap in ai: The need for measuring practitioners' attitudes and perceptions. IEEE Technol. Soc. Mag. (ISSN: 1937-416X) 43 (2), 88–91. http://dx.doi.org/10.1109/mts.2024.3392280.
Said, Nadia, Potinteu, Andreea E., Brich, Irina, Buder, Jürgen, Schumm, Hanna, Huff, Markus, 2023. An artificial intelligence perspective: How knowledge and confidence shape risk and benefit perception. Comput. Hum. Behav. (ISSN: 0747-5632) 149, 107855. http://dx.doi.org/10.1016/j.chb.2023.107855, URL https://www.sciencedirect.com/science/article/pii/S0747563223002066.
Sanguinetti, Pablo, Palomo, Bella, 2024. An alien in the newsroom: Ai anxiety in european and american newspapers. Soc. Sci. (ISSN: 2076-0760) 13 (11), 608. http://dx.doi.org/10.3390/socsci13110608.
Sebastian, Clapp, 2025. Defence and Artificial Intelligence. European Parliamentary Research Service, URL https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2025)769580.
Shin, Yeji, Kim, Seoyong, Kim, Sohee, 2022. Searching for new human behavior model in the climate change age: Analyzing the impact of risk perception and government factors on intention–action consistency in particulate matter mitigation. Int. J. Environ. Res. Public Heal. (ISSN: 1660-4601) 19 (17), 11068. http://dx.doi.org/10.3390/ijerph191711068.
Shinners, Lucy, Grace, Sandra, Smith, Stuart, Stephens, Alexandre, Aggar, Christina, 2022. Exploring healthcare professionals' perceptions of artificial intelligence: Piloting the shinners artificial intelligence perception tool. Digit. Heal. 8, 20552076221078110.
Shneiderman, Ben, 2021. Human-centered ai. Issues in Science and Technology, vol. 37, (2), pp. 56–61.
Slovic, Paul, Fischhoff, Baruch, Lichtenstein, Sarah, 1979. Rating the risks. Environ.: Sci. Policy Sustain. Dev. (ISSN: 1939-9154) 21 (3), 14–39. http://dx.doi.org/10.1080/00139157.1979.9933091.
Slovic, Paul, Fischhoff, Baruch, Lichtenstein, Sarah, 1986. The Psychometric Study of Risk Perception. Springer US, ISBN: 9781461321033, pp. 3–24. http://dx.doi.org/10.1007/978-1-4613-2103-3_1.
Slovic, Paul, Flynn, James, Mertz, C.K., Poumadère, Marc, Mays, Claire, 2000. Nuclear Power and the Public: A Comparative Study of Risk Perception in France and the United States. Springer US, ISBN: 9781475748918, pp. 55–102. http://dx.doi.org/10.1007/978-1-4757-4891-8_2.
Spiel, Katta, Haimson, Oliver L., Lottridge, Danielle, 2019. How to do better with gender on surveys: a guide for hci researchers. Interactions (ISSN: 1558-3449) 26 (4), 62–65. http://dx.doi.org/10.1145/3338283.
Statista, 2022. Artificial intelligence (ai) worldwide – statistics & facts. https://www.statista.com/topics/3104/artificial-intelligence-ai-worldwide/dossier-chapter1. (Last Accessed 28 November 2022).
Surden, Harry, 2019. Artificial intelligence and law: An overview. Georg. State Univ. Law Rev. 35, 19–22.
Tahiru, Fati, 2021. Ai in education: A systematic literature review. J. Cases Inf. Technol. (JCIT) 23 (1), 1–20.
Valdez, André.Calero, Heine, Moreen, Franke, Thomas, Jochems, Nicole, Jetter, Hans-Christian, Schrills, Tim, 2024. The european commitment to human-centered technology: the integral role of hci in the eu ai act's success. I- Com (ISSN: 2196-6826) 23 (2), 249–261. http://dx.doi.org/10.1515/icom-2024-0014.
Van Noorden, Richard, Perkel, Jeffrey M., 2023. Ai and science: what 1 600researchers think. Nature 621 (7980), 672–675.
Veenman, Sietske A., 2013. Futures studies and uncertainty in public policy: A case study on the ageing population in the netherlands. Futures (ISSN: 0016-3287) 53, 42–52. http://dx.doi.org/10.1016/j.futures.2013.08.008.
Verdurme, Annelies, Viaene, Jacques, 2003. Consumer beliefs and attitude towards genetically modified food: Basis for segmentation and implications for communication. Agribusiness (ISSN: 1520-6297) 19 (1), 91–113. http://dx.doi.org/10.1002/agr.10045.
Weber, Max, 1904. Die objektivität sozialwissenschaftlicher und sozialpolitischer erkenntnis [the objectivity of social scientific and social policy knowledge]. Arch. Für Sozialwissenschaft Und Sozialpolitik 19 (1), 22–87. http://dx.doi.org/10.1007/978-3-476-05142-4_65.
Winter, Joos de, Dodou, Dimitra, Eisma, Yke Bauke, 2024. Personality and acceptance as predictors of ChatGPT use. Discov. Psychol. 4 (1), http://dx.doi.org/10.1007/s44202-024-00161-2.
Witte, Kim, Allen, Mike, 2000. A meta-analysis of fear appeals: Implications for effective public health campaigns. Heal. Educ. Behav. (ISSN: 1552-6127) 27 (5), 591–615. http://dx.doi.org/10.1177/109019810002700506.
Wong, Jody Chin Sing, Yang, Janet Zheng, 2021. Comparative risk: Dread and unknown characteristics of the covid-19 pandemic versus covid-19 vaccines. Risk Anal. (ISSN: 1539-6924) 42 (10), 2214–2230. http://dx.doi.org/10.1111/risa.13852.
Yigitcanlar, Tan, Degirmenci, Kenan, Inkinen, Tommi, 2022. Drivers behind the public perception of artificial intelligence: insights from major australian cities. AI & SOCIETY (ISSN: 1435-5655) 39 (3), 833–853. http://dx.doi.org/10.1007/s00146-022-01566-0.
Young, Erin, Wajcman, Judy, Sprejer, Laila, 2023. Mind the gender gap: Inequalities in the emergent professions of artificial intelligence (ai) and data science. New Technol. Work Employ. (ISSN: 1468-005X) 146, 391–414. http://dx.doi.org/10.1111/ntwe.12278.

Dr. Philipp Brauner, is senior researcher at the Chair of Communication Science at RWTH Aachen University, specializing in human–computer interaction, AI perception, and decision-making in socio-technical systems. His research focuses on how people interact with and perceive digital media, including artificial intelligence, emphasizing cognitive biases, risk assessment, and technology acceptance. He has contributed to interdisciplinary projects bridging psychology, computer science, and policy.

Felix Glawe, M.Sc., is researcher at the Chair of Communication Science at RWTH Aachen University, specializing in technology acceptance, AI perception, and human–computer and human–robot interaction. His research focuses on risk–benefit assessments, trust in cyber–physical systems, and societal implications of emerging technologies. He has contributed to interdisciplinary projects examining AI ethics, digital transformation, and user-centered design.

Luca Liehner, M.Sc., is researcher at the Chair of Communication Science at RWTH Aachen University, specializing in human–technology interaction, AI perception, and

decision-making processes. His research focuses on how cognitive and social factors shape public attitudes towards artificial intelligence, with an emphasis on risk assessment, trust, and societal acceptance. He has contributed to interdisciplinary studies on AI governance, ethical technology adoption, and user experience design.

Luisa Vervier, Ph.D., is researcher at the Chair of Communication Science at RWTH Aachen University, specializing in public perception of emerging technologies in the context of the digital transformation of healthcare and production. Her research focuses on how individuals and societies evaluate digital media, with an emphasis on decision-making, trust, and societal acceptance. She has contributed to interdisciplinary projects exploring the intersection human–computer interaction and policy implications.

Prof. Dr. Martina Ziefle is professor of Communication Science at RWTH Aachen University and a leading expert in human–technology interaction, risk perception, and AI governance. Her research explores how individual and societal factors shape the acceptance of emerging technologies, with a strong focus on AI, digital health, and mobility systems. She has led numerous interdisciplinary research projects and has published widely on technology acceptance, user experience, and public perception of innovation.