

BETWEEN ENTHUSIASM AND SCEPTICISM: A MATRIX APPROACH TO ATTITUDES TOWARDS ARTIFICIAL INTELLIGENCE

Malgorzata SUCHACKA

Prof. dr. hab. University of Silesia in Katowice

e-mail: malgorzata.suchacka@us.edu.pl

Abstract

The aim of the analysis is to present an original conceptual model for interpreting attitudes towards artificial intelligence, presented in the form of a two-dimensional matrix. The proposed matrix is based on the distinction between two relatively autonomous dimensions: affective (from lack of trust to full trust) and epistemic (from weak to in-depth narration resulting from knowledge of AI). Such an approach is justified in classical theories of attitudes (Rosenberg and Hovland) and in research on the perception of technological risk (Slovic), which emphasize the importance of both the emotional and cognitive components in shaping the assessment of technology. As a result of the intersection of the two dimensions, four ideal-typical models of approach to AI were distinguished: dogmatic enthusiasm, critical thinking, anxious thinking model and scepticism. This model captures the complexity and ambivalence of attitudes, going beyond one-dimensional shots based on a simple "positive-negative" continuum. The reflection is theoretical in nature and focuses on the methodological justification and analytical potential of the proposed matrix. At the same time, the model serves as the basis for current international empirical research in European Union countries, including in-depth interviews (IDIs), survey research, and focus group interviews (FGIs). The analysis primarily aims to approximate the model's structure and to indicate its applications in the analysis of social representations of artificial intelligence.

Keywords: *artificial intelligence (AI), attitudes towards technology, affective and epistemic dimension, conceptual models, social perception of technology*

1. Introduction

Modern societies are in a phase that the sociology of development refers to as the "great acceleration". This process, driven by radical digitalisation, is not just a technological change, but a profound transformation of social, economic and cultural structures. (Asimov, 2018). Artificial intelligence (AI) has become the focal point of this metamorphosis. While earlier stages of technical development (such as mechanization or automation of production lines) replaced human physical strength, AI is entering spheres previously reserved exclusively for human intellect, emotions and decision-making processes. (Bahishti, 2017).

From the perspective of digital sociology, artificial intelligence has ceased to be merely a set of algorithms confined to IT laboratories. (Castells, 2010). It has become a "social actor" (in the spirit of the actor-network theory) that co-shapes our interactions, influences consumer choices, manages the labor market, and even redefines notions of truth and authenticity. (Latour, 2005, Wojewoda, 2020). The dynamics of this development is unprecedented – technologies that a decade ago were considered a speculative vision of the future, are now integrated into everyday life through voice assistants, algorithms personalizing content in social media or medical decision support systems. (Suchacka et al., 2024). This pace means that the interpretive framework that individuals use to understand the world around them is constantly destabilized.

The use of artificial intelligence algorithms is being analyzed in various spheres of social life. The workplace seems to be a natural area where work efficiency is linked to the use of this technology. It's no surprise that this area is the subject of social research, for example, measuring employee attitudes toward artificial intelligence in the workplace. (Shao et al., 2024).

Despite the ubiquity of AI in public discourse, the sociological approaches to the study of social attitudes towards this technology seem insufficient. An analysis of the literature on the subject and the results of public opinion polls reveals a significant research gap stemming from excessive simplification of the research topic. (Romele, 2024, Stein et al., 2024). Traditional ways of analyzing attitudes are most often based on a one-dimensional continuum, stretched between two poles: technophobia (fear of technology) and technophilia (uncritical enthusiasm). (Slovic, 2000, Montag et al., 2024). In this approach, the individual's attitude is reduced to a simple vector: "positive – negative". Such an approach, while useful for quick opinion polls, is cognitively sterile in the context of theoretical sociology. It does not allow us to understand the phenomenon of technological ambivalence – a situation in which the same individual may simultaneously feel fascinated by the possibilities of AI and deeply fearful of their privacy or economic security. (Fromm, 2014).

A critical analysis of the achievements to date indicates that most studies ignore the interaction between what we feel (the affective component) and what we know about the mechanisms of action of technology (the epistemic component). It is often mistakenly assumed that increased knowledge automatically leads to greater acceptance of technology. Meanwhile, digital sociology shows that in-depth

algorithmic awareness can increase scepticism, while a lack of knowledge and high trust can lead to a dangerous form of digital dogmatism. (Anderson, 2008). Therefore, there is a lack of a tool that would allow us to classify attitudes not only through the prism of "liking" technology, but through the prism of awareness of its functioning.

The main goal of this article is to fill this gap by presenting and justifying the author's matrix model of interpreting attitudes towards artificial intelligence. The proposed model analyses attitudes towards AI in two independent dimensions: the level of trust in the technology and the level of knowledge about its real capabilities and limitations. The combination of these axes allows us to distinguish four types of societal attitudes towards AI and better grasp the complexity of contemporary ideas about AI.

The central research problem of the article focuses on the question: how does the relationship between the level of knowledge and the level of trust differentiate the types of attitudes towards artificial intelligence and what consequences does this have for the inclusiveness of the digital society? In connection with this outlined problem, the following main thesis was put forward: attitudes towards artificial intelligence are not a constant feature of an individual's character or worldview, but the result of a dynamic game between the affective component (trust) and the epistemic component (knowledge). A one-dimensional analysis of technology acceptance masks key societal differences: while knowledge can act as a rationalizing factor for fear (leading to constructive scepticism), the lack of it with high trust leads to vulnerability to algorithmic errors and manipulation.

In this analysis, trust is not treated only as a psychological category, but as the foundation of social stability in conditions of uncertainty. In a world where AI makes decisions about granting credit, diagnosing cancer, or selecting job candidates, trust becomes a "complexity reducer." However, trust without an epistemic foundation (knowledge) is blind trust, which in the sociology of development is considered a barrier to authentic modernization. Knowledge about AI, on the other hand, has an emancipatory function. It allows the individual to move from the role of a passive recipient of technology to the role of a conscious actor who can critically assess the credibility of an algorithmic result. The assumption was made that the key to understanding the social reception of AI is not the question of "whether people accept AI", but "on what basis they build their trust or resistance".

The proposed analysis is discursive in nature and may constitute a theoretical foundation for a broader discussion on the need for new digital education and social policy, which, instead of promoting the uncritical adoption of technology, will seek to shape attitudes based on "critical trust" – a state in which knowledge and emotions work together in the process of taming the digital future.

2. Theoretical foundations of the analysis of attitudes towards artificial intelligence: From classical models to the risk society

Understanding the social reception of AI requires going beyond ad hoc journalistic commentary and embedding the analysis in well-established sociological theories and social psychology. The attitude towards technology is not an isolated phenomenon; It is the result of culture-rooted values, individual cognitive experiences, and emotional responses to the uncertainty that innovation brings. (Rogers, 2003).

The foundation for reflections on the nature of attitudes remains the classical three-component concept, formulated by Milton Rosenberg and Carl Hovland. (1960). Although this model was developed in the middle of the twentieth century, its applicability to research on new technologies remains extremely high, provided it is properly reinterpreted. Rosenberg and Hovland assumed that attitude consists of three interdependent elements: affective, cognitive, and behavioral. In the context of artificial intelligence, the first two components are of particular importance, which are the axes of the matrix proposed in this article:

1. Affective component: refers to the sphere of feelings, emotions and value evaluations. In the case of AI, we rarely deal with emotional neutrality. This technology evokes strong reactions – from fascination (techno-optimism) to existential anxiety (techno-pessimism). These emotions often precede factual knowledge, serving as the primary filter through which individuals process information about new autonomous systems.
2. Cognitive (epistemic) component includes the knowledge, beliefs and arguments that the individual has. This is where our ideas about how an algorithm works, what its limitations are, and what real benefits or threats it brings. (Jonas, 1984),

In the classic approach, high consistency between these components was assumed (if I know that something is safe, I feel confident in it). However, modern digital sociology notices the phenomenon of technological dissonance. An individual may have extensive knowledge of the benefits of AI (cognition) and, at the same time, feel an instinctive fear of losing privacy or autonomy (affect). This disconnect is crucial to understanding why traditional information campaigns (focused solely on education) often fail to change social attitudes. (Habermas, 1997).

Moving on to the sociology of technology, it is impossible to omit Paul Slovic's contribution to the study of risk perception. (2000). Slovic revolutionized thinking about technology by demonstrating that risk is not an objective mathematical parameter (probability times effect), but a social and psychological construct. Slovic introduced the concept of "affect heuristic", which perfectly explains attitudes towards AI. According to the researcher, people assess the risk of a given technology not because of statistical data but on the emotional association that a given object evokes in them. In the case of artificial intelligence, risk perception is determined by two key factors:

- Dread risk: AI is feared because it is perceived as a technology with power unchecked range, capable of catastrophic consequences for the entire species (narratives about the "peculiarity" or the rebellion of machines).
- Unknown risk: AI is new, intangible, and difficult for the average user to understand.

According to Slovic, the less understandable the technology is (low epistemic component), the more its risk assessment is based on emotions (high affective component). This explains why societal attitudes towards AI are so susceptible to media narratives and pop culture – in an information vacuum, imagination and emotions become the main sources of evaluation of technology. (Bostrom, 2014 Méndez-Suárez, et al. 2024).

Referring to Paul Slovic's affect heuristic, attention should be paid to the phenomenon of stigmatization of technology. In the sociology of risk perception, stigmatization occurs when a given technical object is associated with strongly negative events in the public opinion, which leads to a permanent designation of it as "dangerous" or "immoral". In the case of AI, the key trigger for stigmatization is algorithm error. The specificity of social attitudes lies in the fact that a mistake made by a machine is judged much more harshly than an identical mistake made by a human. (Suchacka, et al., 2021). When an autonomous car is involved in an accident, the public does not see it as a single traffic incident but as evidence of the fundamental failure of the entire technology. This phenomenon is based on two mechanisms:

1. Cognitive accessibility: spectacular AI failures are attractive in the media, making them more memorable and shaping the affective component of attitude (anxiety).
2. Loss of systemic trust: Because AI is seen as a precision tool, every mistake strikes at the very core of its legitimacy – the promise of objectivity and errorlessness.

The stigmatization of technology is a factor that rapidly changes individuals' enthusiastic attitude toward AI toward anxious thinking, often without changing the level of actual knowledge about the technology.

To fully understand attitudes, we need to ask: what exactly is AI in society's eyes? AI is often defined as a "black box". This term, popularized by Frank Pasquale, m.in, refers to systems whose internal decision-making processes are opaque to humans. (2015). From a sociological perspective, the "black box" creates an asymmetry of power. The attitude towards such a defined object must be based either on delegated trust (I believe the creators of the system that it acts ethically) or on systemic scepticism. The metaphor of the "black box" finds its solution in the growing importance of the postulates of algorithmic transparency. From a sociological perspective, the opacity of AI is not only a technical problem but, above all, an ethical and political one. If an individual does not understand why the algorithm made a particular decision, they cannot fully trust the system; instead, trust is based on coercion rather than conscious choice. In response to this crisis of trust, the Explainable AI (XAI) trend was born. The postulate of transparency aims to:

- Democratization of knowledge: translating hermetic mathematical language into human-understandable justifications (increasing the epistemic component).
- Building "critical trust": Instead of uncritical faith, XAI proposes trust-based verifiability. A user who understands the machine's decision-making premises ceases to feel manipulated. (Pasquale, 2015).

However, from a digital sociology perspective, transparency is a double-edged sword. On the one hand, it can reduce anxiety (affect), but on the other hand, by exposing the simplifications used by algorithms, it can lead to an increase in substantive scepticism. This confirms the article's central thesis: the increase in knowledge (the epistemic dimension) is not synonymous with trouble-free acceptance. It can even lead to a more conscious rejection of technology, which in our matrix we define as mature, substantive scepticism.

By analyzing attitudes towards AI and the way this concept is understood in society, it is also possible to adopt the perspective of Bruno Latour's actor-network theory (ANT), in which artificial intelligence is treated as a *non-human actor*. This means that AI is not just a tool (like a hammer), but an entity that interacts, imposes certain behaviors on us, and co-creates social reality. (Latour, 2005). Defining AI as an actor changes the nature of attitudes: we no longer just evaluate the "usefulness of the device", but we begin to evaluate the "intentions" and "morality" of the system, which leads to the anthropomorphization of AI.

In the theoretical outline of the selected concepts, it can be noted that attitudes towards AI crystallize at the junction of two great social processes: rationalization (the search for epistemic knowledge) and uncertainty management (affective regulation). The modern sociology of development suggests that the traditional model of trust, based on the authority of institutions, is being eroded in favor of trust based on digital competence. (Campa, 2015). This means that the role of "AI knowledge" is no longer the domain of engineers and has become a key cultural capital for individuals. Without a basic understanding of what machine learning is, the user is left in a state of "anxious ignorance" or "naïve enthusiasm." Trust (affect) and knowledge (episteme) are therefore not separate variables. As Anthony Giddens noted in Modernity Theory, trust is a "bet" against expert systems we cannot fully control. (2010). In the case of AI, this bet is becoming increasingly risky. The matrix proposed later in the article aims to capture how the different configurations of this "bet" (trust) and "security" (knowledge) create a unique map of attitudes in modern society.

The introduction of artificial intelligence into the bloodstream of modern institutions cannot be analysed in isolation from Ulrich Beck's concept of the risk society. (2002). In this approach, AI appears not only as a triumph of the human intellect, but above all as another, complex "side effect" of modernization processes. According to Beck, the late modernity phase is characterized by the fact that the threats produced by technical systems escape the control of the institutions that brought them to life. Artificial intelligence is a paradigmatic example of risk understood in this way. As a product of radical modernization, AI generates threats that are:

- Invisible: algorithmic bias (bias) or automatic discrimination processes take place in the realm of micro-calculation, beyond the reach of human vision.
- Global: A failure of an AI-based operating system or a bug in the global language model can have immediate effects on millions of users on different continents.
- Irreversible: once an algorithmic decision is made (e.g. in credit or recruitment rating systems) it becomes a social fact that is difficult to challenge without deep technical knowledge. (Zuboff, 2020).

In this context, societal attitudes towards AI are a form of response to "manufactured uncertainty". The individual in a society of risk is no longer afraid of nature, but of his own creations. (Fukuyama, 2004).

To sum up, the theoretical landscape of attitudes towards AI spans the technological optimism of modernism and the anxious experience of the risk society. Artificial intelligence, as a "black box" immersed in social networks, requires us to redefine subjectivity. Social attitudes are not just "for" or "against" opinions, but a complicated adaptation strategy. Individuals are trying to balance between the benefits of algorithmization and the risk of losing control, the stigmatization of mistakes and the fight for transparency of processes that increasingly determine their biography. It is these configurations of knowledge and emotions that shape modern digital societies.

3. Methodology of the theoretical analysis – justification

The foundation of this article is the belief that progress in the sociology of technology is not only made through the collection of new empirical data, but above all through the improvement of conceptual tools that allow these data to be organized. (Nowak, 2007). The proposed analysis is theoretical and debatable, a type of research referred to in the methodological literature as conceptual research. This methodology involves deconstructing existing theories and creatively synthesizing them to create a new interpretative framework.

The basic analytical instrument used in the model-building process is the Weberian category of the ideal type (*Idealtypus*). According to Max Weber's view, the ideal type is not a description of reality or its "average" reflection, but a mental construction created by one-sided emphasis on one or more points of view. (1949). As Weber pointed out, the ideal type serves as a "mental image" that allows for the systematic ordering of the chaos of social phenomena. The use of this methodology in the context of attitudes towards artificial intelligence is justified for three reasons:

1. Comparative function: ideal types allow us to measure the degree to which specific, empirical social attitudes approach or deviate from defined poles (affect and episteme).
2. Simplifying complexity: In a world where attitudes towards technology are fluid and multidimensional, the creation of clear, logically coherent categories enables precise scientific communication and the formulation of research hypotheses.

3. Analytical: This approach allows you to go beyond statistical description (how many people trust AI) to understand the meaning (Verstehen) that individuals give to their interactions with technology.

The choice of this methodological path emphasizes that the article does not aspire to be a statistical research report but is an attempt to provide an "understanding matrix" – a tool that other researchers will be able to use to interpret empirical results.

The proposed theoretical model was developed based on unobtrusive research. It is a research approach in which the scientist does not interact directly with the studied individuals but analyzes the products of human intellectual activity. (Przegalińska & Oksanowicz 2020). In this case, the research process is based on two pillars:

1. Critical analysis of secondary sources: this process included a systematic review of sociological and psychological literature and public opinion research reports on AI reception. Elements of content analysis and discourse analysis were used here, which allowed for the abstraction of repetitive thought patterns (e.g., the dichotomy of trust and fear).
2. Abduction as a model of reasoning: the process of arriving at the matrix model was abductive in nature. Unlike induction (from detail to general) and deduction (from general to detail), abduction consists in the search for the best possible explanation for observed phenomena. By analysing contradictions in social attitudes (e.g. the situation of people who have knowledge but still show fear), they were looking for a theoretical structure that would best integrate these apparent inconsistencies.

In theoretical sociology, articles with a review-and-discussion profile serve as a "methodological compass". The importance of such an analysis, detached from ad hoc quantitative measurements, stems from the need to continually redefine concepts in the face of dynamic social change. As Stefan Nowak notes, social science cannot exist without "ordering the language of research". (2007).

The introduction of a new intellectual proposal in the form of a matrix is therefore an activity in the field of sociological theorizing. It is not just an academic exercise, but an essential stage of the scientific process that:

- it legitimizes new perspectives: it allows for the inclusion of previously marginalized categories in the discourse on AI (e.g. the epistemic dimension as a counterbalance to emotions).
- creates interdisciplinary bridges: it combines the achievements of classical sociology (Weber) with modern digital sociology and cognitive science.
- stimulates debate: the discussion article is intended to provoke the scientific community to verify the proposed ideal types in diverse cultural and national contexts.

To sum up, the adopted methodology emphasizes the logical rigor of the theoretical construction. Instead of asking "how many people think in a given way?", we ask here: "according to what logical schemes is the social perception of technology organized?". Such an approach allows the creation of a model that is durable and

resistant to temporary fluctuations in social mood, providing a solid basis for future operationalisations.

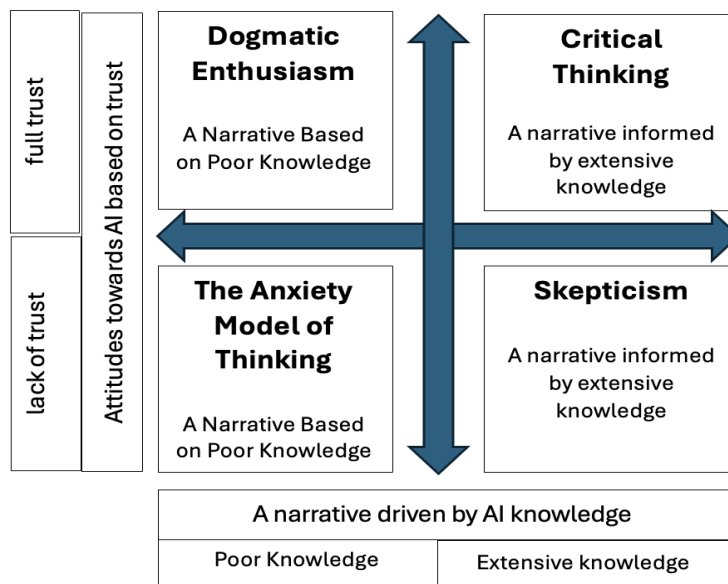
4. Author's theoretical proposal: a matrix of attitudes towards AI

Contemporary digital sociology is challenged to conceptualize the dynamic relationships between the individual and artificial intelligence (AI) systems. The proposed Matrix of Attitudes towards AI is an analytical tool for classifying social reactions. This model, operating on Weber's ideal types, allows us to understand how cognitive competence correlates with emotional and pragmatic attitudes towards algorithms.

The proposed model moves away from a linear perception of attitudes in favor of a two-dimensional system, in which each axis represents an autonomous dimension of the AI's social perception:

1. Affective Dimension: focuses on the level of trust in technology – from paralyzing distrust to full, almost religious belief in the infallibility of autonomous systems.
2. Epistemic dimension: refers to the depth of narrative that an individual builds around AI, based on knowledge of its real capabilities, limitations, and mechanisms (e.g., machine learning, neural networks).

Fig.1. Matrix of attitudes towards AI - conceptual model



Source: Own elaboration - internal documentation of the research team

The intersection of these two axes yields four ideal types: dogmatic enthusiasm, critical thinking, the model of anxious thinking, and scepticism. In this context, societal attitudes towards AI are a form of response to "manufactured uncertainty". The individual in a society of risk is no longer afraid of nature, but of his own creations. (Beck, 2002). Understanding this mechanism is crucial for this matrix – the "anxious" or "sceptical" attitude is not an irrational prejudice, but often a rational response to the awareness that expert systems cannot guarantee full safety.

The following are detailed characteristics of the four quadrants of the matrix, reflecting the main narratives at work in current public and private discourse.

1. Dogmatic enthusiasm: defines an attitude of full trust with a low level of substantive knowledge about the functioning of AI mechanisms. It is a form of uncritical acceptance of technology, often driven by fascination with its "façade" capabilities, without understanding the processes taking place within the so-called "black box". In this model, AI is seen as a kind of modern magic – a versatile, infallible and inherently objective solution. From a sociological perspective, this attitude is associated with technological determinism and unreflective optimism. Individuals in this quadrant often fall victim to technology marketing, taking manufacturers' promises as fait accompli. Lack of technical knowledge means that they are unable to see risks such as algorithmic errors (hallucinations), data bias (bias) or issues related to privacy violations. Trust here is quasi-religious (hence the term "dogmatic"), in which the authority of technology replaces critical judgment. An example would be an office worker who relies entirely on the output of a language model (e.g., ChatGPT) to create financial reports, without verifying the sources or the correctness of the calculations. Its confidence is due to the fact that the tool works quickly and gives the impression of intelligence, although the user has no idea about the statistical nature of the model's prediction of words and its ability to generate completely fabricated data.
2. Critical thinking: this is a quadrant representing a high level of knowledge combined with full trust (or conscious openness). This is the most desirable attitude from the perspective of digital inclusion and the individual's agency. Trust in this model is not "blind", but "earned" – it is based on an understanding of the limitations of technology and the ability to use it safely. In this case, we are dealing with high digital capital. Not only can the unit operate AI tools, but it also understands the logic behind how they work: from machine learning to neural network architecture. This makes trust selective – the user knows when the algorithm is reliable and when it can fail. This is a proactive approach in which AI is treated as an advanced tool to support the human intellect, rather than a substitute for it. Sociology refers to this as reflective technological mediation. An example is a programmer or data scientist who uses AI to optimize code. He knows the architecture of the model, knows what datasets he could have been trained on, and consciously uses "prompt engineering" techniques to minimize the risk of errors. Their trust in the system grows with each successful iteration, but it is always

backed up by substantive verification of the result. Such a person can critically assess the ethics of a given solution before implementing it in operation.

3. Anxiety thinking model: characterized by a lack of trust with a simultaneous lack of reliable knowledge about AI. It is a defensive attitude, often resulting from fear of the unknown and negative media messages (pop-culture visions of machine rebellion). The narrative in this quadrant is dominated by anxiety about losing control, privacy, or work, but these fears are not grounded in technical realities but in a general sense of ontological threat. This phenomenon can be interpreted through the prism of neo-Luddism or resistance to technological change. The lack of understanding of AI's mechanisms makes every innovation seem like a threat to human subjectivity. In digital sociology, there is talk of "mental exclusion", where the barrier of fear prevents the acquisition of competencies, which in turn deepens the lack of trust – a vicious circle of ignorance and fear is created. This attitude often fuels moral panics about artificial intelligence. An example would be an elderly person or a traditional-sector worker who refuses to use any algorithm-based services (even simple recommendation systems or voice assistants), believing that "AI is spying on them" or that "it will soon take over the world." The lack of knowledge about how privacy filters or data encryption work means that every interaction with technology arouses paralyzing resistance, based on visions straight from science fiction movies, and not on real technical parameters.
4. Scepticism: this is an attitude that combines deep substantive knowledge with a lack of trust (or far-reaching reserve) in current AI systems. Unlike the anxiety model, this scepticism is not emotional but grounded on hard evidence, knowledge of systemic limitations and errors in data architecture, and awareness of the political and economic determinants of technology development. This is an attitude often represented by researchers of technology ethics, critical sociologists, and engineers. The scepticism here stems from the knowledge of how easily algorithms can replicate human biases (racism, sexism) or how they are used for surveillance and social manipulation by large corporations (surveillance capitalism). Lack of trust is a political and ethical choice here – the individual knows how the system works, and that is why he does not trust it, seeing it as a tool for deepening social inequalities. An example is a cybersecurity expert or AI ethicist who consciously resigns from using facial recognition systems or algorithmic creditworthiness assessment. His decision is not due to fear, but to knowledge of the imperfections of the teaching sets (e.g. lower effectiveness of recognizing people of skin color other than white) and the lack of transparency of decision-making processes in "black box" algorithms. It is a conscious resistance based on competence.

The matrix of attitudes towards AI shows that the relationship between humans and technology is not linear. The key conclusion of the model is that

knowledge alone does not guarantee trust (as shown by the example of sceptics), and trust without knowledge alone is risky (dogmatic enthusiasm). From the point of view of digital sociology, the biggest challenge for education and social policy is to shift individuals from the quadrants of "poor knowledge" (dogmatic enthusiasm and anxious thinking model) towards the quadrants of "great knowledge". It is only at the level of high competences that it is possible to have a reliable social dialogue about whether we, as a society, want to trust a given technology (critical thinking) or whether we should maintain a substantive reserve towards it (scepticism). This model serves as the foundation for further empirical research on the level of "AI literacy" across various social groups.

5. Discussion - the analytical potential of the model

The proposed matrix of attitudes towards artificial intelligence (AI) is an attempt to theoretically structure a phenomenon that has often been reduced to simplified dichotomies in previous social research. The model's main analytical potential lies in its ability to overcome technological ambivalence. Traditional approaches, based on a linear continuum of "acceptance-rejection", do not allow for capturing intermediate states in which fear coexists with fascination, and substantive knowledge becomes a barrier to trust.

The use of the matrix model allows the sociologist to deconstruct the attitudes of people who are "afraid, but use" technology. Within the matrix, these attitudes are no longer perceived as inconsistent; they become the logical result of the tension between affect and episteme. For example, an individual who shows a high level of trust while having low knowledge (the quadrant of dogmatic enthusiasm) is not simply an "optimist" but a social actor susceptible to technological manipulation and algorithmic errors. On the other hand, a sceptic, who has extensive substantive knowledge, may reject an AI system not because of irrational fear, but through a substantive assessment of its ethical or technical limitations. Such a change in perspective is crucial for analyzing AI social representations, enabling the mapping of public discourses as dynamic shifts across individual fields of the matrix.

The proposed model should be properly operationalized in practice – preferably in qualitative research. The model's analytical potential is most clearly revealed in its implementation within an empirical framework, as exemplified by the workshop "Humans and AI in the World of Tomorrow".¹ This paper shows how theoretical ideal types can be used to stimulate critical thinking in study groups (e.g.

¹ Research project: *"Artificial intelligence and society – towards the coexistence of humans and algorithms"* conducted at the University of Silesia in Katowice (Poland) as part of an international interdisciplinary team consisting of: leader - Małgorzata Suchacka, Anita Pollak, Angelika Pabian, Mirosław Skorek, Susana Amalia De Juana Espinosa (Universitat d'Alicante). The research is carried out from the Dean's own funds of the Faculty of Social Sciences University of Silesia in Katowice, Poland.

students). As part of the proposed workshop activities, participants are asked to identify the dominant ways of thinking about AI through the prism of emotions, arguments and "areas of blindness" characteristic of each quadrant. This approach allows for:

- Visualization of attitudes: by mapping human-AI relationships in specific spheres of life (work, education, relationships), the subjects themselves self-identify within the matrix, which allows the researcher to observe the processes of legitimization of trust.
- Extracting ethical tensions: pinpointing "moments of tension or conflict" in everyday life with AI helps capture the boundary between critical thinking and the anxiety model.
- Analysis of epistemic narratives: auxiliary questions about the degree of control over the individual or the objectivity of algorithms directly examine the component of knowledge (episteme) and its impact on building narratives about technology.

The implementation of the model in international surveys (IDI, FGI, Survey) enables comparisons of intercultural attitudes. It can be hypothesized that in societies with a higher level of digitalization, attitudes on the right side of the matrix (knowledge-based) will dominate, while in societies in the early adoption phase of AI, the discourse will be stretched between dogmatic enthusiasm and a paralyzing fear of the unknown.

An extremely important aspect of the model is its prognostic value. It allows us to predict the direction of evolution of social attitudes under the influence of specific interventions, such as educational campaigns. An increase in the epistemic component does not automatically lead to an increase in trust (acceptance). On the contrary, education can move individuals from the field of anxiety to the field of substantive scepticism, which, from the point of view of digital sociology, is an emancipatory process that builds a more conscious and responsible information society. This model supports the thesis that "the technological future is not given, but co-created". By identifying groups with an anxiety attitude, public policy can precisely target mechanisms to build "critical trust" rather than relying on uncritical promotion of technology, which only leads to the growth of dogmatism.

Despite its wide potential, the proposed matrix has some theoretical limitations that should be considered in the research process:

1. Category static: although the model suggests the possibility of movement between quadrants, it is a static image. In fact, attitudes can be situational – the same person can be sceptical about AI in medicine while being a dogmatic enthusiast of algorithms that personalize entertainment.
2. Simplification of the knowledge component: the "poor-high knowledge" scale is highly subjective and difficult to objectify in quantitative research (the difference between declarative knowledge and actual technical competence).

3. The influence of structural factors: the model focuses on the individual dimension (affect-episteme), thereby underestimating the impact of macrostructural social inequalities (cultural and economic capital), which determine access to knowledge and shape trust in institutions implementing AI.
4. Technological dynamics: Given the pace of AI development, the definition of what constitutes "big knowledge" is constantly being devalued, requiring ongoing updates to the model's epistemic framework.

To sum up, despite the limitations outlined above, the matrix provides an interesting basis for in-depth sociological reflection. It allows us to move from simple statistical descriptions to an understanding analysis of the meanings that modern man gives to his coexistence with intelligent machines. It is a tool not only for the study of reality but also for its conscious shaping, by promoting attitudes grounded in knowledge and reflective trust.

6. Conclusion

The considerations presented here aim to provide a synthetic approach to the issue of social attitudes towards artificial intelligence, perceived not only as a technological artefact but, above all, as a key actor in contemporary modernisation processes. The work on the article allowed us to move from the broad context of digital sociology and the risk society, through rigorous methodological foundations, to the formulation of an original analytical tool in the form of a two-dimensional matrix of attitudes.

The final reflection focuses on three basic levels that define the cognitive value of the analysis: the theoretical redefinition of attitudes, the methodological significance of ideal types, and the practical diagnostic potential of the model.

In the theoretical dimension, the main achievement of the analysis is to demonstrate that traditional, one-dimensional approaches to social attitudes are insufficient in the face of the complexity of algorithmic systems. It has been shown that the relationship between trust and knowledge is autonomous and dynamic, enabling a more complete understanding of technological ambivalence. Embedding the considerations in the classic model of Rosenberg and Hovland, enriched by the perspectives of Paul Slovic and Ulrich Beck, enabled the presentation of artificial intelligence as a social object that generates specific forms of "manufactured uncertainty". The redefinition of attitude as the resultant of the affective and epistemic components makes it possible to move away from sterile disputes between techno-optimism and technophobia, in favour of a multifaceted study of the social framework for interpreting technology.

From a methodological perspective, the article legitimizes the return to Weber's concept of ideal types as an effective heuristic tool in the sociology of technology. The construction of a matrix based on four types – from dogmatic enthusiasm to substantive scepticism – is a proposal to sort out the conceptual chaos accompanying the debate on AI. This approach has been shown to have a high explanatory value, allowing for precise mapping of public discourses and

identification of mechanisms of stigmatization of technology. The use of non-reactive methods and the abductive model of inference allowed for the creation of a theoretical framework that not only describes the existing reality but also serves as an instrument for interpreting dynamically changing phenomena in the digital sphere.

The analytical dimension of the article emphasizes the importance of knowledge as an emancipatory factor, which does not have to lead to uncritical acceptance of technology, but enables the building of subjectivity in relation to the algorithm. The analysis of the model's potential, supported by examples of operationalization in research workshops, shows the practical usefulness of the matrix in diagnosing competence gaps and designing digital education activities. This model allows predicting shifts in social attitudes, indicating that the transition from fear to knowledge-based scepticism is a desirable process from the perspective of the maturity of the information society. At the same time, a critical analysis of the matrix's limitations opens the field for further research on the situationality of attitudes and the influence of power structures on access to epistemic capital.

The presented considerations lead to the conclusion that understanding the social reception of AI requires moving away from simple opinion measurements in favor of studying deep narrative and emotional structures. The proposed matrix model, by combining theoretical rigor with analytical flexibility, constitutes a research proposal adequate to the challenges of late modernity. This summary does not close the discussion but rather invites further verification of the model across diverse cultural and institutional contexts, which seems necessary in the process of consciously co-creating the digital future. The developed matrix thus becomes not only a tool for description, but also a form of social reflection on the limits of human autonomy in the world of algorithms.

References

- Anderson, C. (2008). The end of theory: the data deluge makes the scientific method obsolete, *Wired Magazine*, Vol. 16 No. 7, pp. 16-07.
- Asimov, I. (2018). *Robots and Empire*. Harper Collins Publishers. London.
- Bahishti A.A. (2017). *Humanoid Robots and Human Society* – Vol. 1, Issue 1, pp. 60-63.
- Beck, U. (2002). *The Risk Society: On the Way to a Different Modernity*, Scholar Publishing House.
- Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, Oxford.
- Campa, R. (2015). *Humans and Automata. A Social Study of Robotics*. Berlin, Germany: Peter Lang Verlag. <https://www.peterlang.com/document/1049639>
- Castells, M. (2010). *Network society*. PWN Scientific Publishing House.
- Fromm, E. (2014). *Escape from Freedom*, Reader, Warsaw
- Fukuyama, F. (2004). *The end of man. Consequences of the biotechnological Revolution*, Znak Publishing House, Cracow.
- Giddens, A. (2010). *Modernity and Identity: "I" and Society in the Late Modern Era*, PWN.
- Habermas, J. (1997). *Modernism — an unfinished project*, In: Postmodernism. Anthology of translations, ed. R. Nycz, Cracow.
- Jonas, H. (1984). The Imperative of Responsibility. In: *Search of an Ethics for the Technological Age*, University of Chicago Press, Chicago.
- Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford University Press.
- Méndez-Suárez, M., Delbello, L., de Vega de Unceta, A., Ortega Larrea, A. L. (2024). Factors Affecting Consumers' Attitudes Towards Artificial Intelligence. *Journal of Promotion Management*, 30(7), 1141–1158. <https://doi.org/10.1080/10496491.2024.2367203>
- Montag, C., Ali, R. Davis, K.L. (2024). Affective neuroscience theory and attitudes towards artificial intelligence. *AI&Soc* 40, pp.167–174. <https://doi.org/10.1007/s00146-023-01841-8>
- Nowak, S. (2007). *Methodology of sociological research*. PWN.
- Pasquale, F. (2015). *The Black Box Society: The secret algorithms that control money and information*. Harvard University Press.
- Przegalińska A., Oksanowicz P. (2020). *Artificial intelligence. Inhuman, superhuman. The phenomenon of the world of new technologies*. Social Publishing Institute Znak.
- Rogers E.M., (2003). *Diffusion of innovations*, New York: Free Press.
- Romele A. (2024). *Digital Habitus. A Critique of the Imaginaries of Artificial Intelligence*, Routledge,
- Rosenberg, M. J., Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. In: M. J. Rosenberg, C. I. Hovland (eds.), *Attitude*

- Organization and Change: An Analysis of Consistency among Attitude Components* (pp. 1–14). Yale University Press.
- Shao, C., Nah, S., Makady, H., McNealy, J. (2024). Understanding User Attitudes Towards AI-Enabled Technologies: An Integrated Model of Self-Efficacy, TAM, and AI Ethics. *International Journal of Human–Computer Interaction*, 41(5), pp. 3053–3065.
<https://doi.org/10.1080/10447318.2024.2331858>
- Slovic, P. (2000). *The Perception of Risk* (1st ed.). Routledge.
<https://doi.org/10.4324/9781315661773>
- Stein, JP., Messingschlager, T., Gnamb, T. et al. Attitudes towards AI: measurement and associations with personality. *Sci Rep* 14, 2909 (2024).
<https://doi.org/10.1038/s41598-024-53335-2>
- Suchacka M., Muster R., Wojewoda M. (2021). Human and Machine Creativity: Social and Ethical Aspects of the Development of Artificial Intelligence, *Creativity Studies*, Volume 14, Issue 2: pp. 430–44:
<https://doi.org/10.3846/cs.2021.14316>
- Suchacka M., Wojewoda, M., Nawrat, Z., Domurat, A., Pabian, A.M., (2024). Technical, social, psychological and ethical aspects of human-robot collaboration: a case study of the 'Robin Heart' medical robot, *Journal of Modern Science*, 60/2024, Issue No: 6 Page Range: pp. 369-391
<https://doi.org/10.13166/jms/197008>
- Weber, M. (1949). *The Methodology of the Social Sciences* (E. A. Shils & H. A. Finch, eds.). The Free Press.
- Wojewoda M. (2020). Narratives about Cyborgization in the Context of Technoevolution, *Logos_i_ethos_2020*(52), pp. 11–32,
<http://dx.doi.org/10.15633/lie.3575>
- Zuboff, S. (2020). *The Age of Surveillance Capitalism: The Struggle for the Future of Humanity on the New Frontier of Power*, Wydawnictwo Zysk i S-ka.

