SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2025
ORGANIZATION AND MANAGEMENT SERIES NO. 233

ANTHROPOMORPHIC DESIGN: ETHICAL DILEMMAS

Grażyna OSIKA

Silesian University of Technology; grazyna.osika@polsl.pl, ORCID: 0000-0002-8729-1001

Purpose: The main objective of this article is to describe the ethical aspects of anthropomorphic design in terms of the possibility of excluding unethical practices.
Design/methodology/approach: A number of psychological machisms are used in the design of UX/UI for digital solutions, one of which is anthropomorphism, i.e., an automated cognitive response; this fact can raise ethical questions related to the loss of autonomy by users. Hence, there is a need for analyses dedicated to the ethics of the scope of their influence. Two research methods were used in the deliberations, a bibliometric study and a conceptual analysis. The bibliometric analysis helped to establish the importance of this issue of design anthropomorphism, while the conceptual analysis helped to develop a quasi-model for the ethical assessment of its impact.
Findings: This paper first establishes the growing body of research dedicated to the use of the anthropomorphisation mechanism in design activities. Second, due to the automatic nature of the cognitive response that is anthropomorphism, attention was drawn to the possibility of ethical dilemmas related to design. Thirdly, within the framework of applied ethics, a quasi-model of moral assessment of design impact has been developed, which may be helpful in identifying so-called dark design patterns. The research work undertaken may be helpful in eliminating unethical forms of design.
Research limitations/implications: These considerations are only a conceptual sketch and, as such, require in-depth theoretical analyses, which will then be subjected to empirical verification, e.g. piloting, e.g. how often the anthropomorphisation mechanism is used in design or what user awareness is related to the use of this mechanism, whether they accept the use of automated responses in UX/UI design, etc.
Practical implications: The research conducted in this article has a very practical dimension, as the theoretical analysis undertaken was aimed at creating tools to identify cases of unethical use of the anthropomorphism mechanism, ultimately this can be used to identify so-called dark design patterns and eliminate their negative impact on society.
Social implications: The article is conceptual in nature, which, in line with the research method used, allows the development of a theoretical reflection (primary research), in this case concerning the use of an automatic cognitive mechanism for design activities, This analysis can be useful to identify unethical uses of this mechanism and contribute to building autonomy towards its impact.
Originality/value: The most significant achievement of the present reflections is to develop a quasi-model to subject UX/UI projects that use the mechanism of anthropomorphism in their design activities to an ethical evaluation.

Keywords: anthropomorphism, design, anthropomorphic design, applied ethics, ethics of influence; ‘dark patterns’.
Category of the paper: Conceptual paper.

1. Introduction

Joseph Wiezenbaum, creator of the first chatbot, ELIZA, in *Computer Power and Human Reason*. *For Judgment to Calculation,* published in 1976, cites as one of the main motives for writing this book very simple recognitions that require understanding: "I was startled to see how quickly and how very deeply people convering with DOCTOR became emotionally involved with the computer and how unequivocally they anthropomorphized it. Once my secretary, who had watched me work on the program for many months and therefore surely knew it to be merely a computer program, started conversing with it. After only a few interchange with it, she asked me to leave the room. […] I was promptly bombarded with accusation that what I proposed amounted to spying on people’s most intimate thoughts: clear evidence that people were conversing with the computer as if it were a person who could appropriately and usefully addressed in intimate terms" (1976, pp. 6-7).

Similar comments were made by Sherry Turkle, who was among those who had the opportunity to test the first chatbot and watch how students, knowing that the programme knew nothing and understood nothing, wanted to talk to it and be alone with it, implicitly entrusting it with their secrets. In her research work, S. Turkle pointed out the tendency to see machines as "human" agents, "if they can, for example, make eye contact, track our movements and make friendly gestures" (2013, p. 29). In her view, we must have programmed some evolutionary ‘buttons’ activated during contact with robots that caused us to think and behave towards them as if we were another human being. We now know that anthropomorphisation is a cognitive predisposition involving the attribution of human nature to non-human objects including technological ones; it concerns the processes of perception and meaning-making at the most basic level, so related phenomena remain outside human consciousness.

This tendency was already recognised in antiquity, one of its motivations being related to the situation of reducing uncertainty, helping to cope with its effects. But while for J. Wiezenbaum this fact was a human concern, over time, this mechanism began to be used in design processes to guide human reactions and behaviour, this applies to the UX/UI (User Experience/User Interface) of digital solutions. It seems that this situation is not ethically neutral, as symptoms of a loss of decision-making autonomy can be recognised in the technologically biased behaviour of users, hence the need to problematise these issues.

From this point of view, it is crucial to understand the mechanism of anthropomorphisation and its principles of operation, i.e., to find the evolutionary ‘button’ mentioned by Turkle, as well as to identify the design areas that make use of this mechanism. The research work

undertaken will provide a closer look at the specifics of this cognitive predisposition, so that it can then be analysed from an ethical point of view. In this section, the perspective of applied ethics will be used, thanks to which it will become possible to isolate possible ethical dilemmas. The paper assumes that an awareness of the ethical dilemmas involved in the use of anthropomorphisation mechanisms in the design of digital technology will ensure a more balanced impact of these dilemmas and thus contribute to maintaining a sense of autonomy towards technology. We need to know the pitfalls so that, as users, we can consciously build mental models of adequate responses to the suggestive potential of digital instrumentation.

2. Methods

The paper will use two research methods: a bibliometric method and a conceptual analysis. Bibliometric analysis is a quantitative and statistical method of examining the literature; it involves the use of a variety of data on scientific publications and citations to assess the performance of scientific activity, monitor the development of science, identify trends and impact in a field (OECD Glossary of Statistical Terms). It includes a quantitative assessment of scientific output (e.g. number of articles, books, chapters); a citation study, an indicator of the impact and importance of a given work is studied; a study of the impact of citations of a given scientist or scientific unit; a study of the average number of citations of articles published in a given journal over a given period; a study of networks between scientists and institutions; identification of key topics and their links in a given field; visualisation of links between publications, authors, institutions, or keywords, helping to uncover trends and research gaps; measures scientific activity in the digital world (e.g. mentions in social media, blogs, scientific websites), complementing traditional indicators. It is used to assess the scientific output of individual researchers, research units, institutions, and even entire countries. It is used for evaluation, recruitment, or awarding of prizes and grants; identification of research trends, helping to track the development of scientific fields, the emergence of new research areas, and the disappearance of others; identification of the most influential researchers in a given field; and is particularly useful for the creation of systematic reviews and meta-analyses, as it allows a quantitative approach to the analysis of large amounts of data. In the present discussion, the number of publications with the two key words "anthropomorphism" and "design" will be analysed, helping to identify the number of publications dedicated to the issue of design anthropomorphism, the type of publications describing the phenomenon, and the research area. This will allow us to trace the development of the research with an indication of the main threads of the analysis.

The second method used for consideration will be one of the oldest scientific methods, i.e. conceptual analysis (Furner, 2004; Gilson, Goldberg, 2015; Stuart, 2015; Dickson et al., 2018), whereby further theoretical analyses are conducted on the basis of already existing knowledge. In this method, using the most classical research methods, such as deductive reasoning, initial assumptions, well-grounded in theory, are made, and from these initial assumptions, conclusions are formulated that extend knowledge in the field or allow new hypotheses to be put forward, which are then subjected to verification in empirical research. This type of research is typical of so-called basic research. Conceptual analysis is used to combine theories, adapt theories to new developments, categorise, establish logical relationships between phenomena, and build theoretical models (Jakkolla, 2020). Given the theoretical sophistication of the issues discussed in the considerations, this method seems appropriate. According to the conceptual analytical steps adopted, the study will carry out the following tasks:

- Defining basic concepts, describing the initial theoretical assumptions, in this case, defining terms such as ‘anthropomorphism’, ‘anthropomorphic design’, ‘applied ethics’, ‘ethics of influence’.
- Establishing the relationships between the different concepts, in this study, the authors answer the question: In which areas of design is the mechanism of anthropomorphism used?
- Conclusions – in this study, an answer to the question: What ethical dilemmas might be generated by the use of the mechanism of anthropomorphisation in design?

The theoretical knowledge obtained in this way can form the basis for empirical research, as hypotheses developed that should be verified in the course of further research work (Osika, 2024, pp. 457-458).

3. Results

3.1. Bibliometrics

According to the methodological assumptions described above, it is initially necessary to establish to what extent the proposed topic is currently a relevant research topic. Below are the results of a search in the SCOPUS database from the Science Direct website, covering terms such as: "anthropomorphism", "design". Data were collected on 30.06.2025. The number of publications by year, type of publication, and scientific discipline was taken into account. The results obtained are presented sequentially in Figure 1 (by years), Figure 2 (by type of publication), Figure 3 (by scientific field).

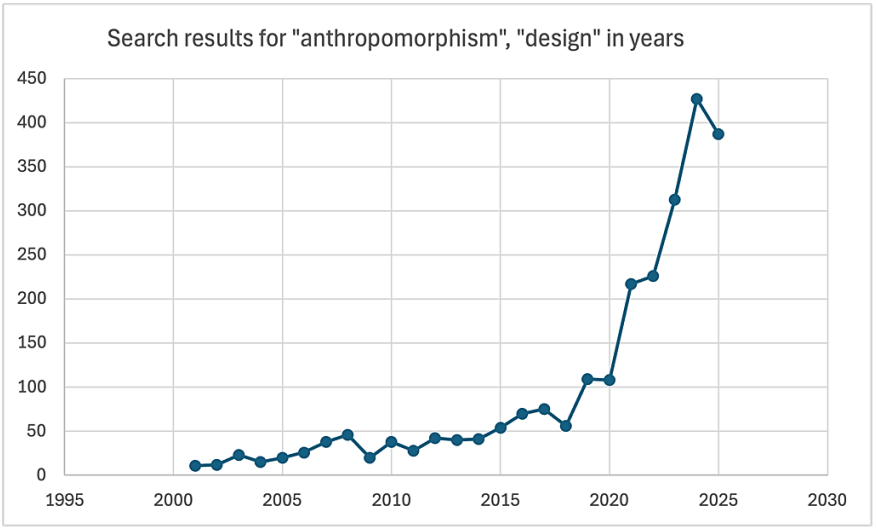

Figure 1. SCOPUS-based Science Direct – number of published articles search for "anthropomorphism" and "design" in years.

Source: Own elaboration.

Considering the number of publications appearing in a given year for searches of the terms "anthropomorphism" and "design", it can be seen that since 2020 their number has been steadily increasing. This may indicate a constant demand for knowledge in this area motivated by design considerations, i.e. that anthropomorphism as a cognitive tendency is used in UI/UX.

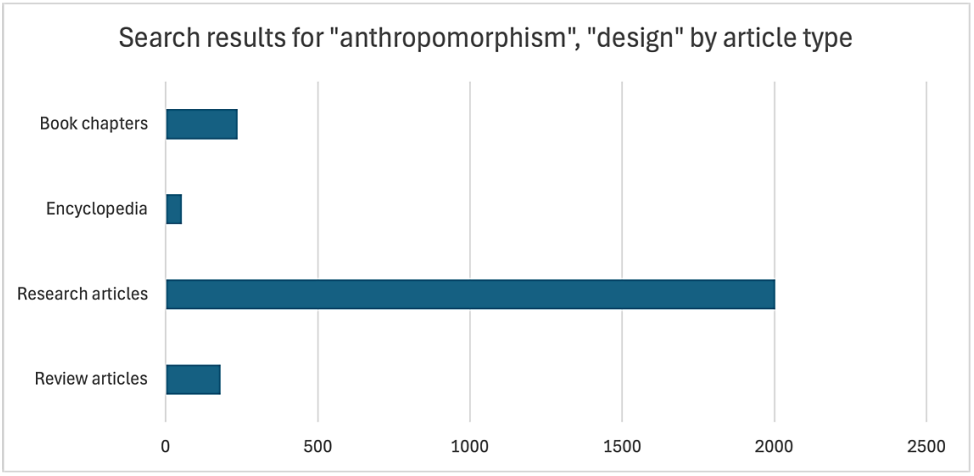

Figure 2. SCOPUS-based Science Direct – number of published articles search for "anthropomorphism" and "design" by article type.

Source: Own elaboration.

The above-mentioned trend can also be seen in the type of research conducted. As shown in the chart above, research publications are the most numerous, which means that we are currently at the stage of gathering knowledge.

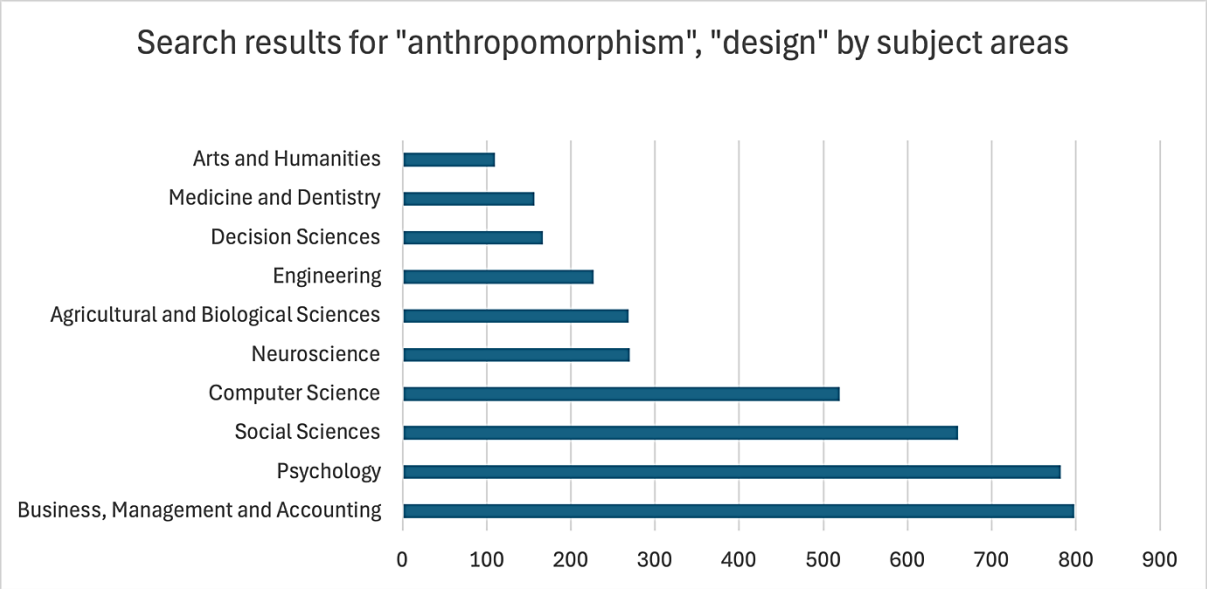

Figure 3. SCOPUS-based Science Direct – number of published articles search for "anthropomorphism" and "design" by subject areas.

Source: Own elaboration.

Similar conclusions can be drawn with regard to the identified research areas, i.e. psychology, computer science, business, management and accounting, social science – psychology and social sciences concern research into the phenomenon of anthropomorphism itself, while the other two are likely to be related to the use of findings.

The data obtained from the bibliometric study allow several conclusions to be drawn, indicating the existence of certain trends related to scientific analyses of issues concerning anthropomorphism in design. First, an increase in interest in this issue has clearly been discernible in the past five years. The decrease in the number of publications in 2025 is the result of a lack of complete data; there are still a few months to go before the end of the year, and the current number of publications rather indicates a continuation of the trend. The increasing interest of researchers usually indicates that the subject matter is problematised, i.e. that it is becoming the focus of attention, as an issue of importance, and this in turn is the result of increasingly perceived gaps in current knowledge. Secondly, the type of publications that predominate are research articles, which may indicate a pragmatic tendency in the approach to results, i.e. the use of research results in project activities. This may also be evidenced by the research areas within which the publications are produced mentioning business, management and accounting, psychology, social science, and computer science.

The bibliometric analyses obtained in the course of the research indicate that the scientific work carried out is intended to serve design activities enabling greater use of knowledge about this largely automated cognitive tendency, bypassing conscious responses, from an ethical point of view this trend may be questionable.

3.2. Conceptual analysis

3.2.1. Definition of anthropomorphism and anthropomorphic design

Anthropomorphism, a concept derived from the Greek words anthropos meaning "human" and morphe referring to "shape" or "form", we should identify with all forms of attributing life to the inanimate (Epley et al., 2007, p. 864). This mechanism refers to human properties, attributes or mental states real or imagined projected onto non-human "subjects" and objects; these attributes include experience, metacognition, intentions, emotional states, but also behavioural traits (ibid.). Cecilia Roselli, Leonardo Lapomarda and Eduardo Datteri, in their article How culture modulates anthropomorphism in Human-Robot Interaction: A review (2025), remind us that ‘tendency to "humanize"’ everything was a subject of philosophical analysis in ancient Greece, where Xenophanes was the first to use the term ‘anthropomorphism"t’o illustrate how Greek gods were invariably light-skinned and blue-eyed, whereas African gods were invariably dark-skinned and dark-eyed"(Roselli Cecilia Roselli, Leonardo Lapomarda and Eduardo Datteri, in their article How culture modulates anthropomorphism in Human-Robot Interaction: A).

In detail, several key features can be identified that help reveal the essence of anthropomorphism. This process can be understood as the cognitive predisposition to attribute human nature or characteristics of human uniqueness to non-human technological, mental, or natural objects (Ruijten, 2015, pp. 11-13; Roselli et al., 2025, p. 1; Kowalik, 2015; Szpunar, 2023). "If we encounter an object whose action we do not understand, cannot explain its action or behaviour, we superimpose matrices familiar to us from the human social world, i.e. human matrices, onto the situation. In other words, anthropomorphisation provides us with an intelligible and useful vocabulary through which that which is incomprehensible and alien somehow becomes closer and tamer, because it contains a human element" (Szpunar, 2023, p. 28).

One of the most important and influential theories explaining why people attribute human characteristics (thoughts, intentions, emotions) to non-human entities is considered to be the three-factor theory of anthropomorphisation, formulated by Nicholas Epley, Adam Waytz and John Cacioppo (2007). Identifies three main psychological factors that determine the propensity to anthropomorphise. This is elicited agent knowledge, a factor that refers to whether people possess knowledge or cognitive schemas that can be applied to understand and predict the behaviour of a given non-human agent. In other words, the more an object appears to have human-like characteristics, such as eyes, a face, a voice, the ability to communicate, and movement at a human pace, the easier it is to anthropomorphise it. The second factor is effectance motivation, which refers to the human need to control and understand the environment. When people encounter non-human entities whose behaviour is unclear, unpredictable or difficult to explain, they may anthropomorphise these entities to give them meaning and a sense of control. Attributing human intentions, thoughts or emotions to them

can make their behaviour seem more understandable and predictable, and thus easier to control. (Epley et al., 2007; Waytz et al., 2010). The last factor relates to sociality motivation and refers to the human need for belonging, social connection, and avoidance of loneliness. When people feel a lack of social connection or loneliness, they may anthropomorphise non-human entities (e.g., pets, electronic devices) to satisfy their social needs. They then treat them as 'friends', 'companions", or 'confidants'. This mechanism helps to reduce feelings of isolation and provides a substitute for social interaction (Epley et al., 2007; Epley et al., 2008; Szpunar, 2023).

Anthropomorphism as a cognitive strategy can be implicit and spontaneous or intentional and reflexive (Ruijten, 2015, p. 13), it can be thought of psychologically, i.e. when people perceive something as 'human', but also technologically, when, for design reasons, a tool is consciously given human characteristics (Ruijten, 2015, pp. 15-16), we then speak of anthropomorphic design (Fischer 2021). This recognition directly relates to the *Computers Are Social Actors* (CASA) model, which is one of the most important theoretical contributions to the field of human-machine communication. This model suggests that humans apply the same heuristics that are used in human-to-human interactions to non-human artificial agents (e.g. computers, chatbots and robots) and therefore interact with them as if they were humans (Nass, Siminoff, Steuer, 1994; Nass, Moon, 2000; Roselli et al., 2025).). In other words, individuals unthinkingly, i.e. using specific heuristics, apply social rules and expectations to machines, and it should be anticipated that this also applies to any digital solution. Two dimensions seem to be key in these diagnoses; on the one hand, there is the automatic way of reacting, combined with the social character of this reaction, both of which are considered as design cues to trigger a psychological response in humans. In this sense, Kerstin Fischer proposes to consider design anthropomorphism as the conscious shaping of a response aimed at triggering anthropomorphising behaviour in humans, i.e. to obtain in humans an observable behavior in contact with technological objects, analogous to the one they manifest towards each other (Ficher, 2021, pp. 4-2). Anthropomorphisation in the design of robot, for example, is conceived as triggering psychological mechanisms that lead people to attribute human-like characteristics to robots, thus ensuring smooth and successful interactions between them (Roselli et al., 2025).).

It is now possible to speak of a significant body of work, both scientific and applied, realising the extent of the use of design anthropomorphism. Andreas Janson (2023, p. 3) has reviewed such research and applications, here are some examples he describes:

- 2009: Lingyn Qiu and Izak Benbasat – a study of the impact of anthropomorphic design in terms of modality and performance on the effectiveness of a recommendation agent - conclusions: communication with embodiment and voice (compared to plain text) positively affects 'social presence', as well as trust, perceived friendliness and propensity to use a recommendation agent.

- 2010: Kathleen Keeling et al. – a study of the impact of communication styles on trust and intention to visit a website - findings: task-oriented communication style influences trust, especially when searching for goods/services. Socially oriented communication contributes to trust and increases the intention to visit the website.
- 2011: Clemens F. Köhler et al. – a study of the influence of conversational content and social actor style on new customer adaptation - conclusions: Both functional and social content have a positive impact on customer adaptation, but there must be a balance between the two conversational styles.
- 2014: Tibert Verhagen et al. – investigating anthropomorphic tendencies and communication style as moderators of perceived social presence and personalisation - conclusions: a socially oriented communication style reinforces the positive impact of benevolence and developing competences focused on personalisation also develops the virtual agent's sense of social presence, and social presence and personalisation have a positive impact on user satisfaction.
- 2015: Jennifer Hill et al. – exploring differences in the quality and content of conversations between humans and humans interacting with machines - findings: dialogues between humans were more varied and longer. More messages were used in dialogue between humans and chatbots.
- 2018: Theo Araujo – investigating the effect of name and language style on the activation of anthropomorphic perceptions – conclusions: anthropomorphic perceptions towards a chatbot are activated towards agents with a human name and using an informal language style, and this in turn allows stronger emotional connections to be built with the company represented by the human chatbot.
- 2020: Adam et al. – investigating the relationship of the anthropomorphic elements used in the design and the foot-in-the-door technique on building user compatibility; conclusions: there is a positive relationship between the anthropomorphic perception-evoking elements used and the use of the foot-in-the-door technique on perceived compatibility with the chatbot.
- 2021: Rajat Roy and Vik Naidoo, study of the effect of conversational styles triggering anthropomorphic perceptions on users't'emporal orientation and evaluation of chatbot effectiveness, conclusions: chatbot effectiveness depends on users' temporal orientation and informal conversational style, both parameters indirectly affect the perception of the brand that uses chatbots to communicate.
- 2023: Joohee Kim and Il Im – investigating the effect of a non-human actor's appearance, on the release of anthropomorphic perceptions related to the chatbot's intelligence rating - conclusions: the anthropomorphic response depends on perceptions of the agent's appearance and intelligence rating. Users attribute a greater range of humanity to agents perceived as intelligent but disembodied than to

> ‘intelligent’"‘and"poorly designed’ physical agents; A. Janson – analysing anthropomorphic aspects, personification and communication style as attributes influencing the perception of chatbots. His research shows that the personification of the chatbot, but also the social orientation of the communication, have a clear impact on the perception of the chatbot. Both positively influence the level of recognised social presence, but only communication style directly influences satisfaction with the chatbot. Furthermore, social presence is a fundamental parameter of importance when evaluating a chatbot, with factors such as trust, perception of empathy, and feelings of satisfaction from contact with the chatbot (Jason, 2023, p. 4).

These examples point to the significant ‘social suggestiveness’ of digital technology, which is achieved by design through the use of automatic cognitive and behavioral strategies, depriving users of control over their reactions because they are reflexive in nature, and this fact generates the dilemmas highlighted by Brian. J. Fogg, admonishing "designers must be aware of the ethical implications of designing psychological cues into their products" (Fogg 2003, p. 100). More and more attention is being paid to this design aspect, and the term ‘dark patterns’ has even been coined to describe unethical design practices using, among other things, cognitive automatisms (Luguri, Strahilevitz, 2021; Kho, Seah., 2023; Roffarello, Russis, 2023; Ehsan, Riedl, 2024; Yablonski 2024). The following section of the article is devoted to this issue.

3.2.2. Ethical dilemmas of anthropomorphic design

Intuitively, we recognise that using mechanisms in design that are triggered in users automatically, outside of their control, brings concerns of ethical violations, but having a hunch is not the same as trying to understand the situation and consciously decide what kind of influence we are able to accept. Hence, the need for a little more in-depth reflection. However, this requires the adoption of a conceptual framework that allows it to be carried out according to methodological rules. First of all, it is necessary to choose an adequate research area, and in this case it is worth following a pragmatic criterion, as it is not a question of building theory but of simple indications allowing one to separate the ‘bright patterns’ of design from the ‘dark’ patterns" ones. It seems that, in this case, the most appropriate would be applied ethics, which from its essential "tries to give answers to the practical moral questions we ask in everyday life (Jackson, 2021, p. 1). Secondly, develop moral evaluation criteria specific to the activities in question, taking into account the context, these are ethical criteria related to social impact. Third, analyze the project activities in terms of the criteria developed earlier.

3.2.2.1. Applied ethics

According to Aleksandra Kuzior, "the term applied ethics began to be used only in the 1970s. This trend of practical philosophy grew out of a certain powerlessness in ethics itself. Applied ethics (or rather practical ethics) is an ethical reflection on practical problems and attempts to solve them thanks to the tools used" (Kuzior, 2021, p. 18). In Encyclopedia Brittanica we read "applied ethics, the application of normative ethical theories i.e. philosophical theories regarding criteria for determining what is morally right or wrong,

good or bad—to practical problems". However, applied ethicists are not only an activity to which moral principles are supposed to apply, but also a method because applied ethics "refers to any use of philosophical methods to treat moral problems, practices, and policies in the professions, technology, government, and the like" (Beauchamp, 2003, p. 3).). Regardless of the doubts raised by the very notion of applied ethics, formulated mainly by philosophers (Sójka, 2007, pp. 120-1220), several characteristics can be mentioned. Firstly, it involves all stakeholders, and above all practitioners, for whom ethical dilemmas arise in action. Secondly, it must be subject to updating, adapting to real-life actions, and thirdly, it is mainly concerned with dilemmas that play out socially, helping to establish the rules we as a society want to follow. Fourthly, in its analyses it refers to specific cases in order to be able to formulate more general recommendations on the basis of these, and fifthly it is dialogical, i.e. its findings are worked out by the consensus of the stakeholders involved in the dilemma and their arguments (Sójka, 2007, pp. 122-124). Finally, practical dilemmas provide the context within which normative concepts are formulated (Darwall, 2003, p. 18; Sójka, 2007, p. 124).

In summary, applied ethics, using methods of analysis typical of normative ethics, diagnoses dilemmas and attempts to formulate the principles needed to resolve them in relation to specific areas of action (Beauchamp, 2003; Allhoff, 2011; Attfield, 2022; Meynell, Paron, 2023).

3.2.2.2. Applied ethics of social influence

In line with the tenets of applied ethics, the criteria used for moral judgement are always situated in relation to a particular form of action. Similarly, in the case of exerting social influence, it is necessary to identify a framework that makes judgement possible; in turn, the construction of criteria requires the delineation of reasonably precise denotational areas to refine the understanding of social influence. Lisa Rashotte defines social influence as change in an individual's thoughts, feelings, attitudes, or behaviors that results from interaction with another individual or a group (Rashotte, 2007). In turn Philips G. Zimbardo i Michaele R. Leippe emphasise that the process of social influence involves a person's behaviour that has the effect (or is born out of such an intention) of changing another person's behaviour, feelings or thoughts about the stimulus. [...] There are many techniques of social influence, but they all boil down to relatively few, very basic processes related to the human way of thinking, remembering, feeling and making decisions (Zimbardo, Leippe, 2004, pp. 19-20; Osika, 2005; Osika, 2012). Influence takes place in several distinct interaction systems, but the tool of influence is communication, largely persuasive. It seems that nowadays, limiting influence to shaping information flows is insufficient, as social influence processes are increasingly thought of in terms of experience design (Norman, 1988; Osika, 2019). In this case, the components that determine the influence exerted need to be recognised in a new way. Bernd H. Schmitt lists six strategic modules of experience, these are: sensory arousal; cognitive processes such as learning, evaluation of preferences, appreciation of excellence; emotions; higher (spiritual)

experiences; a sense of agency; and social relationships and the associated sense of belonging and identity (Schmitt, 1999; Boguszewicz-Kreft, 2021).

In relation to design anthropomorphism, this extended spectrum of influence will be of particular importance due to the accumulation of "triggers" for automatic responses that guarantee efficiency. And it is this high range of efficiency that is particularly threatening from an ethical point of view because it creates the possibility of taking control over the user. On the other hand, it also seems unjustified to havetily label any form of design anthropomorphism as unethical, because when a robot designed for human collaboration "betrays human symptoms" designed during its construction in order to facilitate human comfort, we will doubt whether this is a form of non-ethical action.

Dealing with such dilemmas requires establishing the criteria for moral judgement that we can adopt; within applied ethics it is common to situate judgements in relation to human acts (Hartman, 2008; Cichosz, 2008; Osika, 2012). These criteria include, for example, "the intentions and motives of the actor, the moral qualities (virtues) that come to the fore in a given deed, the moral values that are realised by a given deed; the conformity of the deed with the applicable rules of conduct and the commitments undertaken; the conformity of the deed with the conscience" (Hartman, 2008). In the criteria developed, it is also worth taking into account two traditions of evaluating actions, i.e. the intentionalist one, which emphasises the intention of the act, and the consequentialist one, which emphasises the role of the effects of actions. This condition is fulfilled by the set developed by Marcin Cichosz, which includes: the assessment of motives, the assessment of intentions, whereby in the case of motives it is about the cause of the action, and in the case of intentions it is about the intentions accompanying the action; the third criterion is the components after the execution of the action, i.e. the assessment of the relation of the action to social norms, the assessment of the person acting, and also the assessment of the effect (Cichosz, 2008, p. 51).). The indicators introduced by Cichosz build a grid that allows for an ethical analysis of exerting influence (Osika, 2012), as we can juxtapose the moral evaluation of project intentions with the ethical judgement of the actual effects that users are more or less aware of.

3.2.2.3. Antropomorphic design – areas of uncertainty

As the subject of the article is design activities, only these will be subjected to ethical evaluation, given the narrowness in the title regarding areas of influence, the examples cited will focus on the use of the mechanism of anthropomorphism. However, it seems that this analysis can be more general, and, due to the categorisation grid developed, it can be considered as a kind of quasi model for ethical evaluation of project activities. In order to give more clarity, it will be in tabular form. It will use the criteria proposed by Cichosz, framed comparatively as 'ethical activities'/'non-ethical activities'. - this will bring out the ambiguity of using the mechanism of anthropomorphisation during design and the moral dilemmas that arise.

Table 1.
Categorisation grid for ethical design impact assessment

	Before action			After action	
	Motives	**Intentions**	**Social norms**	**Person acting**	**Effects**
Ethical actions	Designers combine the user's needs with their own, e.g. the anthropomorphisation mechanism is incorporated into the design brief, but it is intended to create a sense of security for the user; this in turn allows the company to generate profit by building its brand, and the design group gains a reputation as professionals	When designing a particular product, designers use anthropomorphism to enhance the user's quality of life, e.g. a rehabilitative technological solution that encourages frequent use of the	When designing a given product, designers use anthropomorphism to mimic behavioural norms, e.g. facilitating a relationship with technology by using polite conventions when interacting with a chatbot	When designing a given product, designers use anthropomorphism to make it easier for the user to act with it - Jakob's law - 'people using a digital product or service immediately know how to do it' (Yablonski, 2024, p. 1)	The effect of using the anthropomorphisation mechanism is reduced stress when using a digital device
Unethical actions	Designers use the mechanism of anthropomorphisation during design, but it is intended to trigger automatic responses to cover design deficiencies or make users dependent on the product	Designers, when designing a given product, use anthropomorphism to build trust to force user settings that facilitate data capture	When designing a particular product, designers use anthropomorphism to mimic behavioral norms in order to trigger a behavioral pattern, such as building a strong emotional connection with the company	Designers use anthropomorphism when designing a given product by 'sewing' elements of their stereotyping behavior into a digital solution	The effect of using the anthropomorphisation mechanism during design is to create psychologically harmful behaviors and habits

Source: Own elaboration.

The suggestions in the table above show that the ethical evaluation of design solutions that use an automatically triggered anthropomorphisation mechanism can be carried out against a given solution at the level of motives and intentions accompanying the design, and, as it were, post factum, when the design has been implemented we can monitor the effects of the design at the level of norms, the evaluation of those acting, and the ethical dimension of the effects of using a given digital solution. An important fact emerges from this analysis, namely that the evaluation of design intentions can be verified by an impact evaluation, which may suggest that a reflection integrating both types of evaluation is necessary to be able to fully demonstrate the real impact and the positive or negative sign of this impact.

It is also worth emphasising that in the case of motives and intentions, it is the moral awareness of those designing and implementing a given solution that is important, while in the case of effects, it is the users themselves or certain institutions acting on their behalf who should take control measures and exclude unethical projects from the social space. Attention is already paying more and more attention to the so-called 'dark patterns' (Ehsan, Riedl, 2024; Yablonski, 2024), the identification has educational values for users, as it makes them aware of the mechanisms used to change their attitudes. On the other hand, the creation of lists of such

patterns is a form of moral rectification of the design community, making it possible to speak clearly on the ‘light side of power’, as certain types of solutions are ‘"’stigmatized".

An example of this is the ‘notifications’ feature. The creators of this functionality may have had pure intentions, aiming to provide users with neutral information about changes that could be important to them. Over time, modifications to the strategy of building economic value in social networking services meant that notifications enabled in applications automatically became the unethical tool for exploiting attention that they are today, but only after the fact can we assess the effects. In turn, with regard to anthropomorphism, let us assume that the designer of a gambling application consciously uses human attributes to trigger automatic reactions designed to win the user's favour and make them addicted to it. In this case, our ethical assessment will concern intentions and motives rather than effects. In this sense, the quasi-model of moral assessment is universal because it allows us to refer to both situations, i.e. intentions, motives and effects.

The proposed model for ethical evaluation of design solutions is not perfect; however, its aim was more to start a discussion on the necessity of undertaking this type of reflection in view of the fact that automatic cognitive mechanisms are used in the design of solutions that are now becoming a “natural” part of our living environment. Consequently, their influence, unconscious to us, can significantly direct our behaviour and habits, not necessarily in accordance with our intentions and values. Therefore, it is worth reclaiming the right to make responsible decisions in this regard.

4. Discussion

From the very beginnings of the emergence of digital technology and the emergence of artefacts associated with it, we have observed symptoms of anthropomorphisation made to both devices and programmes (Wiezenbaum, 1976; Turkle, 2013). This is hardly surprising, as the cognitive tendency to humanise the inanimate has been recognised since antiquity in relation to gods, animals, objects, etc. (Epley et al., 2007; Ruijten, 2015; Roselli et al., 2025; Kowalik, 2015). One conception that allows us to understand the workings of this mechanism points to three main psychological factors, it is the triggering of the schema on the basis of human resemblance, in the case of uncertainty anthropomorphisation helps to make sense, and in the case of loneliness it helps to offset ‘social hunger’ (Epley'a, Waytza, Cacioppo, 2007; Epley et al., 2007; Epley, 2008; Szpunar, 2023). Mostly anthropomorphic reactions are automatic in nature and serve as a way for humans to alleviate dissonances of a social nature. One type of anthropomorphism is the so-called technological anthropomorphism in which, for design reasons, a given tool is consciously given human characteristics (Ruijten, 2015; Nass, Siminoff, Steuer, 1994; Nass, Moon, 2000; Roselli et al., 2025). This approach is about the deliberate

shaping of responses to trigger anthropomorphizing behaviour in humans (Ficher, 2021). In these cases, the aim is to ensure smooth and successful interactions with the product, service (Ficher, 2021; Roselli et al., 2025). Many aspects of the impact of design anthropomorphism have already been studied (Janson, 2023; Qiu, Benbasat, 2009; Keeling et al., 2010; Köhler et al., 2011; Verhagen et al., 2014; Hill et al., 2015; Araujo, 2018; Shemekhi et al., 2018; Adam et al., 2020; Roy, Naidoo, 2021; Kim, Im, 2023; Janson, 2023). And it is precisely this fact that provokes ethical considerations as to whether the use of automated responses violates the autonomy of users. The most appropriate area for this kind of reflection seems to be applied ethics, as that field of ethics in which ethical reflection on practical problems is undertaken using methods typical of normative ethics (Darwall, 2003; Beauchamp, 2003; Sójka, 2007; Allhoff, 2011; Kuzior, 2021; Meynell, Paron, 2023). In this reflection, it was necessary to address the essence of influence (Rashotte, 2007; Zimbardo, Leippe, 2004; Osika, 2005) and the ethics of influence (Fogg, 2003; Osika, 2012), understood as the ethics of action (Hartman, 2008; Cichosz, 2008; Osika, 2012). It can be expected that this model will be helpful in delineating so-called 'dark patterns' (Luguri, Strahilevitz, 2021; Ehsan, Riedl, 2024; Yablonski, 2024) capable of restoring control over design processes, through awareness of what mechanisms of automatic perception are used during design, an educational value, and to initiate work on institutional prohibition of unethical design practices.

5. Conclusions

As indicated in the “Introduction”, UX/UI design for digital solutions today uses a number of psychological machisms, which, by design, are intended to respond to user' needs. However, when we are aware that the mechanisms rely on largely automated user responses, an analysis of the ethical aspects of this situation seems obvious. Two research methods, a bibliometric study and a conceptual analysis, were used to carry out such an analysis. The bibliometric analysis has established a growing interest in the issue of anthropomorphism in design reference, so the problem is gaining attention. On the other hand, as part of the conceptual analysis, according to the assigned research tasks:

1. concepts such as anthropomorphism, design anthropomorphism, applied ethics, applied ethics of social influence were defined,
2. the relationships between the key concepts were established, including in which design areas the mechanism of anthropomorphism is used,
3. Through the development of a quasi-model of moral evaluation, the existence of possible ethical dilemmas was indicated, while obtaining a tool to facilitate such evaluation.

Further research highlighted the need to identify so-called "dark design patterns", as suggested by Mustafa Suleyman, co-founder of the pioneering artificial intelligence development companies DeepMind and Inflection AI, in the development of technology we can expect a number of processes to slip through our fingers. Therefore, it is sometimes necessary to take containment measures that are complementary in technical, social, legal dimensions; this containment is to address mechanisms that help monitor the development of technology at each level (Suleyman, 2024, p. 28). This monitoring should also include design processes, but for such monitoring we need a categorisation framework to help understand what we are monitoring. The developed quasi-model for ethical assessment of the use of design anthropomorphism can act as one such tool.

As previously indicated, in accordance with the principles of applied ethics, certain normative assumptions constituting an analytical framework are necessary for the moral evaluation of actions. The creation of these assumptions is conceptual in nature, constituting a theoretical model, but its success is ultimately determined by 'tests' on specific actions. In the case in question, it will fulfil its function if we can use it to analyse any UI/UX solution and assign it the status of an ethical or unethical model. This requires further analytical work.

References

1. Allhoff, F. (2011). What Are Applied Ethics. *Sci Eng Ethics*, *Vol. 17*, pp. 1-19.
2. Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, *Vol. 85*, pp. 183-189.
3. Attfield, R. (2022). *Applied Ethics. An Introduction.* Cambridge: Polity Press.
4. Beauchamp, T.L. (2003). *The Nature of Applied Ethics.* In: R.G. Frey, C.H. Wellman (eds.), *A Companion to Applied Ethics* (pp. 1-16). Malden: Blackwell Publishing.
5. Boguszewska-Kreft, M. (2021). *Marketing doświadczeń. Jak poruszyć zmysły, zaangażować emocje, zdobyć lojalność klientów.* Warszawa: CeDeWu.
6. Cichosz, M. (2008). Znaczenie intencji dla moralnej oceny czynu – stanowisko Daniela Wegnera. *Analiza i Egzystencja*, *Vol. 8,* pp. 49-63.
7. Darwall, S.L. (2003). *Theories of Ethics*. In: R.G. Frey, C.H. Wellman (eds.), *A Companion to Applied Ethics* (pp. 17-37). Malden: Blackwell Publishing.
8. Dickson, A. et al. (2018). Theoretical and Conceptual Framework: Mandatory Ingredients of A Quality Research. *Quality Research, International Journal of Scientific Research*, *Vol. 7*, pp. 438-441.
9. Ehsan, U., Riedl, M. (2024). Explainability pitfalls: Beyond dark patterns in explainable AI. *Patterns, Vol. 5*, pp. 1-8.

10. *Encyclopedia Brittanica*. Retrieved from: https://www.britannica.com/topic/applied-ethics, 30.06.2025.
11. Epley, N., Akalis, S., Waytz, A., Cacioppo, J. (2008). Creating social connection through inferential reproduction: loneliness and perceived agency in gadgets, gods, and hreyhounds: Research article. *Psychological Science, Vol.19(2)*, pp. 114-120.
12. Epley, N., Waytz, A., Cacioppo, J.T. (2007). On Seeing Human: A Three-Factor Theory of Anthropomorphism. *Psychological Review, Vol. 114/4*, pp. 864-886.
13. Fischer, K. (2021). Tracking Anthropomorphizing Behavior in Human-Robot Interaction. *ACM Transactions on Human-Robot Interaction, Vol. 11(1)*, pp. 1-28.
14. Fogg, B.J. (2003). *Persuasive Technology. Using Computers to Change What We Think and Do.* San Francisco: Morgan Kaufmann Publisher.
15. Furner, J. (2004). Conceptual Analysis: A Method for Understanding Information as Evidence, and Evidence as Information. *Archival Science, Vol. 4*, pp. 233-265.
16. Jakkola, E. (2020). Designing conceptual articles: four approaches. *AMS Review, Vol. 10*, pp. 18-26.
17. Janson, A. (2023). How to leverage anthropomorphism for chatbot service interfaces: The interplay of communication. *Computer in Human Behavior, Vol. 149*, pp. 1-17.
18. Kim, J., Im, I. (2023), Anthropomorphic response: Understanding interactions between humans and artificial intelligence agents. *Computers in Human Behavior, Vol. 139(11)*, Retrieved from: https://www.sciencedirect.com/science/article/abs/pii/S0747563222003326, 05.05.2025.
19. Konya-Baumbach, E. et al. (2023). Someone out there?: A study on the social presence of anthropomorphized chatbots. *Computers in Human Behavior, Vol. 139(4)*. Retrieved from: https://www.sciencedirect.com/science/article/abs/pii/S0747563222003338, 05.05.2025.
20. Kuzior, A. (2021). *Applied Ethics*. Lublin: Wydawnictwo Naukowe TYGIEL.
21. Luguri, J., Strahilevitz, L.J. (2021). Shining a Light on Dark Patterns. *Journal of Legal Analisis. Harvard Law School, vol. 13*, pp. 43-109.
22. Meynell, L., Paron, C. (eds.) (2023). *Applied Ethics*. Primer, Bedford: Atlantic Canada: Pressbooks Network.
23. Nass, C. Siminoff, E., Steuer, J. (1994). Computer are Social Actors. *Human Factors in Computing Systems*, pp. 72-78.
24. Nass, C., Moon, Y. (2000). Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues, Vol. 56(1)*, pp. 81-103.
25. Norman, D.A. (1988). *The Psychology of Evertday Things*. New York: Basic Books.
26. Osika, G. (2005). Ogólna charakterystyka procesu perswazji. *Zeszyty Naukowe Politechniki Śląskiej, seria: Zarządzanie, Vol. 140*, pp. 1-16. Retrieved from: https://open.icm.edu.pl/server/api/core/bitstreams/09c4cb4f-49ca-4242-ba83-44674a83d62a/content, 5.07.2025.
27. Osika, G. (2012). Etyczne aspekty wpływu społecznego. In: L. Karczewski, H. Kretek (eds.), *Odpowiedzialny biznes i konsumeryzm wyzwaniem XXI wieku* (pp. 393-407).

Racibórz: Wydawnictwo Państwowej Wyższej Szkoły Zawodowej w Raciborzu. Retrieved from: https://open.icm.edu.pl/items/b4aff157-620f-482b-a7bc-d00d18ea17a8, 5.07.2025.

28. Osika, G. (2019). Design User Experience jako podstawa budowania relacji w sektorze kreatywnym. *Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach, Vol. 388*, pp. 122-133.
29. Qiu, L., Benbasat, I. (2009). Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems. *Journal of Management Information Systems, Vol. 25(4)*, pp. 145-182.
30. Rashotte, L. *Social influence,* Retrieved from: http://www.blackwellpublishing.com/sociology/docs/BEOS_S1413.pdf, 5.07.2025.
31. Roselli, C., Lapomarda, L., Datteri, E. (2025) How culture modulates anthropomorphism in Human-Robot Interaction: A review. *Acta Psychologica, Vol. 255*, pp. 1-12.
32. Roy, R., Naidoo, V. (2021). Enhancing chatbot effectiveness: The role of anthropomorphic conversational styles and time orientation. *Journal of Business Research, Vol. 126(2)*, pp. 23-34.
33. Ruijten, P.A. (2015). *Responses to Human-Like Artificial Agents.* Eindhoven: Technische Universiteit Eindhoven.
34. Schuetzler, R.M. (2020). The impact of chatbot conversational skill on engagement and perceived humanness. *Journal of Management Information Systems, Vol. 37(3)*, pp. 875-900.
35. Seeger, A-M. et al. (2018). Designing anthropomorphic conversational agents: Development and empirical evaluation of a design framework. *ICIS 2018 Proceedings*, pp. 1-17.
36. Shamekhi, A. et al. (2018). Face value? Exploring the effects of embodiment for a group facilitation agent. In: R. Mandryk, M. Hancock, M. Perry, A. Cox (eds.), *Proceedings of the 2018 CHI conference on human factors in computing systems - CHI '18* (pp. 1-13), Montreal QC, New York: ACM Press.
37. Sheehan, B. et al. (2020). Customer service chatbots: Anthropomorphism and adoption. *Journal of Business Research, Vol. 115(2)*, pp. 14-24.
38. Sójka, J. (2007). Idea etyki stosowanej: Pomiędzy wyżynami filozofii a wymogami praktyki. *Annales*, *Etyka w życiu gospodarczym, Vol. 10(1)*, pp. 117-126.
39. Stolpe, K., Hallström, J. (2024). Artificial intelligence literacy for technology education. *Computers and Education Open*, *vol. 6,* pp. 1-8.
40. Stuart, M.T. (2015). Philosophical Conceptual Analysis as an Experimental Method. In: T. Gamerschlag, D. Gerland, R. Osswald, W. Petersen (Eds.), *Meaning, Frames and Conceptual Representation* (pp. 267-292). Dűsseldorf: Dűsseldorf University Press.
41. Suleyman, M., Bhaskar, M. (2024). *Nadchodząca fala. Sztuczna inteligencja, władza i najważniejszy dylemat ludzkości XXI wieku.* Kraków: Szczeliny.

42. Szpunar, M. (2023). Między antropomorfizacją maszyn a technicyzacją człowieka. *Journal of Moder Science, Vol. 3(52)*, pp. 24-38.
43. Weizenbaum, J. (1976). *Computer Power and Human Reason. For Judgment to Calculation.* New York/San Francisco: W.H. Freeman and Company.
44. Yablonsky, J. (2024). *Prawa UX. Jak psychologia pomaga w projektowaniu lepszych produktów i usług.* Gliwice: Helion, O'REILLY.
45. Zimbardo, P.G., Leippe, M.R. (2004). *Psychologia zmiany postaw i wpływu społecznego*. Poznań: Zysk i s-ka.