Artificial Intelligence Review (2025) 58:337
https://doi.org/10.1007/s10462-025-11306-7

AI-mediated healthcare and trust. A trust-construct and trust-factor framework for empirical research

Marcos Alonso[1] · **Aníbal M. Astobiza**[2] · **Ramón Ortega Lozano**[3]

Accepted: 16 June 2025

Abstract
Application of Artificial Intelligence (AI) in healthcare is growing exponentially, and its use is expected to continue expanding in the coming years. However, lack of trust in AI systems remains a significant barrier to their widespread adoption. This article analyzes the problem of trust, its various features and its application in AI-mediated healthcare. We first review the literature on trust and trust in technology to detect which theoretical constructs are essential to trust. We then identify the factors that we consider fundamental for a rich and complex comprehension of trust in AI-mediated healthcare. We finally propose a trust-factor framework that could be used for empirical research on AI-mediated healthcare and its practical implementation.

Keywords Artificial intelligence · Healthcare · Trust · Ethics · Bioethics · Medical ethics

1 Introduction

One of the most profound developments of our time is the widespread adoption of artificial intelligence (AI) across all sectors of society (Maslej et al. 2024). The recent development of generative AI has only accelerated and deepened this repercussion (Sim and Cassel 2024). The impact is especially swift and noticeable in healthcare, where AI has the poten-

✉ Marcos Alonso
marcos.alonso@ucm.es

Aníbal M. Astobiza
amastobiza@ugr.es

Ramón Ortega Lozano
rortegal@comillas.edu

[1] Complutense University of Madrid. Department of Public Health and Maternal-Child Health. Faculty of Medicine., Madrid, Spain

[2] Department of Philosophy I, University of Granada, Granada, Spain

[3] San Juan de Dios Foundation Comillas Pontifical University, Health Sciences Department. San Juan de Dios School of Nursing and Physical Therapy, Madrid, Spain

Published online: 20 August 2025

tial to revolutionize diagnostics, treatment planning, and patient care, boosting efficiency and accuracy (Nittas et al. 2023). Yet, this rapid integration also introduces significant risks and challenges, given the sensitive nature of healthcare.

AI in healthcare has the potential to completely transform patient care and clinical decision-making (Asan et al. 2020, 2). AI technologies in healthcare can help with diagnostic, prognostic, and therapeutic decision support (Xing, Giger and Min, 2021). Notable examples include AI systems for medical imaging in medical specialties such as radiology, dermatology, ophthalmology, cardiology, and oncology. These systems make use of machine learning algorithms to improve the precision of diagnoses and offer individualized therapy suggestions. Research into healthcare AI is primarily driven by the potential to lower the alarmingly high rates of wastefulness and human mistake in medical practice, in addition to the potential elimination of laborious administrative activities and the enhancement of cost-efficiency in medicine (Almyranti et al. 2024). The application of AI in healthcare, however, presents many difficulties.

For example, the implementation of AI in healthcare is fraught with ethical challenges. Vereschak et al. have warned that the use of AI in medicine could contribute to "compromising safety and health of individuals, discrimination, and harming human dignity" (2021, 1). As Sparrow and Hatherley (2019) point out, AI in medicine brings with it threats to privacy and the intensification of surveillance, due to the massive data collection necessary to train AI systems. We also have to deal with the issue of biases, which can seriously harm underprivileged communities. The use of AI in healthcare systems has the potential to redistribute power, favoring computer scientists over conventional doctors. Additionally, excessive reliance on AI can introduce “single points of failure” (A Single Point of Failure -SPOF- is a critical system component whose failure causes the entire system to stop functioning), increasing the system's fragility. The growing presence of AI may dehumanize medicine, eroding the doctor-patient relationship, which is fundamental to care. Last but not least, the “black box” nature of many AI algorithms makes it difficult for healthcare professionals to understand and trust the system’s decisions fully.

In the face of all these dangers and problems, trust seems to have emerged as a precious and desired goal, as recent guidelines seem to defend (Jobin et. al., 2019). The rationale is clear: if we develop trust in these technologies, if we have a trustworthy AI, we will be free of most challenges and its use will be unimpeded. This focus on trust has given rise to the so-called TAI (Trustworthy AI) paradigm (Floridi 2019; Thiebes et al. 2020). This notion of trust and trustworthiness is presented, especially in various recently published guidelines, as a kind of *panacea*, transforming trust and trustworthiness "into an umbrella term for all things in general considered'good'" (Reinhardt 2023, 738). The concept of trust and trustworthiness—whose application to AI is, for some authors, a "notable conceptual misunderstanding" (Hatherley 2020, 3)—thus becomes overloaded and its operability is significantly reduced. Moreover, the guidelines tend to focus on AI developers and providers, without adequately considering how end users might come to trust AI systems (Li et al. 2024, 2).

In this article we analyze the theoretical and practical problem of trust, its definition and its application in healthcare AI systems. In order to identify the main theoretical structures related to trust, we first review the literature on trust and trust in technology. We then list the factors that we believe are essential for a deep understanding of trust in artificial intelligence within the healthcare context. To conclude, we offer a trust-factor framework that may be applied to empirical studies on the application of AI in healthcare. Our main premise is that

identifying the factors that increase, reduce or more generally affect trust in AI-mediated healthcare will allow us to obtain a better understanding of this reality. This improved comprehension will consequently facilitate the design of better empirical approaches to this problem where we can measure and evaluate trust, ultimately serving as a basis for the implementation of trust-enhancing technologies and environments, as well as a foundation for the education of health professionals.

2 AI-mediated trust in healthcare. Theoretical background and some caveats

The study of the concept of trust has a long history in philosophy and academia (McLeod 2020). Luhmann (1968) and Rotter (1967) broadly define trust as the expectation of the reliability of others' promises. This initial definition already implies a basic distinction between what we might call the subjective pole of trust (trustor, that is, the person who trusts), the objective pole of trust (trustee– typically, the person or object who receives trust), and its relational context. Anticipating what will be our view, contrast between trustor and trustee must always be linked to what we will call the relational level, which has been somewhat narrowly thought of in the literature as "context-related factors".

More recently, considering explicitly a technological context, trust has been understood as the willingness to be vulnerable to the actions of another entity, expecting it to act in a beneficial manner (Mayer et al. 1995). From this view, trust is the firm belief or hope in the reliability, truth, or competence of someone or something. However, from the outset, we observe how the anthropological dimension repeatedly infiltrates almost all attempts to define trust. This is because "trust is a fundamental human mechanism that is required to cope with vulnerability, uncertainty, complexity, and ambiguity" (Choung et al. 2022, 10). It is precisely the uncertain situation we find ourselves in today, with the emergence of AI and its use in sensitive areas, that leads us to repeatedly, and perhaps excessively (Reinhardt 2023, 735–736), resort to trust as a safeguard against this situation.

An initial caveat is that trust is not inherently positive. Trust is not good in itself. As Reinhardt (2023, 736) points out, trust can be excessive, inappropriate, or blind, issues that only some guidelines manage to highlight. As Asan and colleagues also explain, "maximizing the user's trust does not necessarily yield the best decisions from a human-AI collaboration" (2020, 4), thus advocating for what they call "optimal trust" (Asan et al. 2020, 4), referring to a justified form of trust that is neither excessive nor deficient. In this vein, a risk not always highlighted in the literature is the possible "overtrust in AI systems" (Li et al. 2024, 9). As Reinhardt concludes, it may be that "in the end what we need is not more trust in AI but rather institutionalized forms of distrust" (2023, 741). Distrust, precisely, can also be valuable and positive -even though we tend to confuse distrust with low levels of trust, which is a mistake (Vereschak et al. 2021, 11)- as this could involve a methodological skepticism preventing catastrophic consequences arising from excessive and unjustified forms of trust. While it is generally true, as Hengstler et al. state, that "trust is an evolving and fragile phenomenon and can be destroyed much more quickly and easily than it can be created" (2016, 106), this does not imply that we should unconsciously and irresponsibly embrace the endeavor of fostering trust in AI in any manner and at any cost.

If we delve onto the specifics, we can see that the literature tends to agree in that there are certain irreducible differences between interpersonal trust and trust in technological systems. Interpersonal trust can be defined as the generalized expectation regarding the reliability of another person's words or promises. Trust in technology is the expectation, based on past interactions, that the actions of a technology are in line with one's expectations and advantageous to oneself (Gefen 2000). Zuboff (1988) explains that trust in a new technology depends on trial-and-error experience, followed by an understanding of how the technology works, and finally, faith in that technology. The evolution from reflections on HCI (Human–Computer Interaction) and HRI (Human–Robot Interaction) to HAI (Human-AI Interaction) (Ueno et al. 2022) has derived in the TAM theoretical framework (technology acceptance model), originally developed by Davis (1989), that has been used to analyze the acceptance of various technologies over the past decades (Venkatesh & Davis 2000; McLean & Osei-Frimpong 2019). Trust in technology in the AI-mediated healthcare context has been already addressed by the works of Calhoun et al. (2019), Kohn et al. (2021) and Gillath et al. (2021).

We already have evidence that AI, used as a healthcare assistant, improves medical assistance (Matheson 2019). However, we must be aware that AI is not just another medical tool. We are not dealing with a stethoscope or a scalpel, the context and relational dimension changes significantly; AI presents the possibility of undermining the practical and epistemic authority of humans (Sim and Cassel 2024). AI could replace humans, or at least fully substitute their actions and healthcare tasks (Choung et al. 2022, 11). Furthermore, AI calls into question the epistemic authority of human doctors (Hatherley 2020, 3). AI has demonstrated capabilities comparable to or even superior to those of human healthcare providers in areas such as medical history analysis, clinical data interpretation (Davenport and Kalakota 2019), image-based diagnosis (Gulshan et al. 2016; Forghani et al. 2019), treatment plan design (Davenport and Kalakota 2019), hospital readmission prediction (Caruana et al. 2015), virtual nursing (Peete et al. 2019), new drug discovery (Ramsundar et al. 2015), and selecting suitable study participants in clinical trials (Beck et al. 2020). Diagnosis, prognosis, and treatment selection, fundamental tasks of the clinician according to Eric Cassell (2004), could be placed in the hands of AI, producing a paradigm shift with consequences difficult to anticipate (Topol 2019).

Trust in AI may differ from interpersonal trust in some decisive respects. For example, due to the lack of intentionality in AI systems. We believe trust is still possible in these cases when one pole of the trust-relationship cannot have intentions, but it is clear that trust is transformed in these situations. Trust in people is commonly defined as the willingness to depend on another due to the characteristics of that other, such as benevolence. However, this concept is less applicable to technology, because the trustee in this case lacks volition and moral agency. As McKnight et al. explain, “trust in people and trust in technology differ in terms of the nature of the object of dependence” (2011, 5), also suggesting that taking into account trustor’s expectations requires us to replace benevolence with helpfulness, competence with functionality, and integrity with reliability. Some authors have distinguished certain concepts or constructs more related to human trust, such as integrity, honesty, or benevolence, from other concepts or constructs more related to trust in systems, such as reliability, functionality, or utility (Lankton et al. 2015).

This can be especially significant in the healthcare field and its relational implications. In this domain, trust has traditionally been considered a crucial element, both intrinsically,

by creating a unique bond between patient and healthcare professional, and instrumentally, by facilitating treatment acceptance and improving outcomes (Hatherley 2020, 1). Some authors specifically refer to the irreducible level of intentions by asserting that reliability is not sufficient to generate trust, as trust requires a belief in the goodwill and correct motivations of the agent (Hatherley 2020, 3). AI systems, lacking will and motivations, could not truly be objects of trust. Trust in humans implies a moral responsibility that, Hatherley argues, cannot be transferred to machines.

Nevertheless, while trust in human-AI relations may be an improper way of speaking, the fact is that the inclusion of AI in healthcare affects the entire healthcare context (Sparrow et al., 2020). It therefore affects trust between doctors and patients (Nundy et al. 2019); trust between doctors, nurses, and essentially all possible combinations of healthcare professionals (Hall et al. 2002); trust regarding healthcare institutions and their administrative staff (Balkrishnan et al. 2004). In all these cases, we believe it may make more sense to speak of "AI-mediated trust" rather than trustworthy AI; although this cannot probably go beyond a theoretical warning, as the discussion around «trustworthiness» is already well established and therefore is not easily avoidable.

3 Trust-related constructs. Subjectual, objectual and relational constructs

An initial distinction can be established, from a phenomenological perspective, between what we might call the subjective pole of trust (trustor), the objective pole of trust (trustee), and its relational context. The difference between trustor and trustee consists in the difference between the person who trusts and the one person or thing in which they trust (Mayer et al. 1995; McKnight et al. 2011, 6). In any trust relationship, the trustor is the one who gives the trust, and the trustee is the one who receives it and must act accordingly to maintain it. However, distinctions between trustor and trustee as different poles of trust should not be overemphasized. Beyond this contrast between trustor and trustee, there can be considered what we might call a relational level, which has been somewhat narrowly thought of in the literature as “context-related factors”. Among these would be perceived uncertainties and benefits, regulations, laws, safeguards, social influence, or certain cultural factors (Li et al. 2024, 8; Baer et al. 2018; Westjohn et al. 2022). As Reinhardt aptly notes regarding technology:

> Technologies are not developed in a societal vacuum, but are interwoven with the fabric of our social and political interactions. They are, as socio-technical systems, embedded in societal and political contexts. This is particularly true with regard to AI technologies and algorithmic decision making; algorithms based on machine learning already shape our lives and social interactions in profound ways (Reinhardt 2023, 740)

This point is crucial, and we must never forget how our conceptions shape and are shaped by our technologies, generating a feedback loop in which it may be difficult or even impossible to distinguish clearly between causes and effects. It’s from this point of view that we

can explore what we will call trust-related constructs, that is, constructs that are usually connected thematically to trust and directly influence trust relationships.

In understanding trust in AI, it's crucial to identify related constructs such as accuracy, explainability, and accountability. We believe it is useful to group them into subjective, objective, and relational aspects. Trust, however, is inherently relational, bridging subjective and objective dimensions. What this classification seeks to do is help our understanding of each construct's role in trust, especially in AI-mediated healthcare. On one hand, subjective constructs -or "subjectual" to bypass the negative connotation of subjective- emphasize user perception, including motivation, past experiences, and individual propensity to trust (Brown et al. 2004). Motivation and prior experiences play pivotal roles; positive interactions with AI foster trust, while negative experiences can breed skepticism. Familiarity with AI also enhances trust (Gefen et al., 2003), highlighting the importance of positive initial interactions. Knowledge about AI, reducing anxiety and increasing trust, leads to the crucial issue of interpretability (Lipton 2018). Interpretative capacity (what could be understood as the subjectual reverse of explainability -even though interpretative capacity and explainability are neither necessarily subjectual or objectual) is also crucial for trust. On the other hand, objectual constructs are grouped when they primarily point to AI's performance. These include capacity (Malle and Ullman 2021), accuracy (London 2019), reliability (London 2019), and robustness (Asan et al. 2020). These attributes enhance trust by ensuring AI's competence and consistency. Explainability is critical (Shin 2021a), but must be balanced, as understanding it as an absolute requirement would be unwarranted (Li et al. 2024).

At any rate, the relational dimension is the most fundamental one, and therefore a phenomenological perspective is demanded. Constructs like responsibility, accountability, anthropomorphism, and value alignment emphasize shared human-AI dynamics. Responsibility involves collaboration between humans and AI, fostering trust through justified actions (Matthias 2004; Leo and Huh 2020). Accountability promotes trust by ensuring that any action taken by AI can be explained and justified, and that the consequences of such actions will be assumed and, if necessary, compensated (Shin and Park 2019). Conversely, "when people think that AI cannot be held accountable, they are less willing to let AI make decisions" (Li et al. 2024, 7). Anthropomorphism, giving AI human-like qualities, can enhance trust by facilitating emotional connections (Cominelli et al. 2021; Waytz et al. 2014; Kim et al. 2018; Shin 2021b). Value alignment ensures AI acts in accordance with users' ethical principles, emerging from a dynamic, bidirectional relationship (Israelsen and Ahmed 2019).

Finally, there are constructs often connected to trust that we think should not appear in the discussion, as they can be misleading. Constructs like controllability, transparency, and security. Castelfranchi and Falcone, in fact, go as far as to state that control and trust are "two opposite notions" (2010, 192). True trust involves managing, not eliminating, uncertainty. Controllability can imply a lack of trust, as "trust implies some (perceived) lack of controllability" (Castelfranchi and Falcone 2010, 82) -although other authors consider controllability can be an alternative to trust (Kieseberg et al. 2023). Transparency, while beneficial for fairness, is not a substitute for trust, which accepts some unknowns. Security, protecting against risks, is essential but not synonymous with trust. Overemphasis on these constructs can empty the concept of trust, as trust inherently involves navigating uncertainties. Thus, understanding trust in AI requires a nuanced approach, recognizing the interplay of subjectual, objectual, and relational constructs.

4 Trust-factors in AI healthcare

The previous sections provide a panoramic view of trust in AI and its associated conceptual constructs. Our aim was to clarify some discussions in the literature, dispelling misconceptions and elucidating the articulation between various categories frequently used when discussing trust and AI. There have been significant precedents of trust-frameworks in the literature. For example, Li and colleagues (2024) propose a three-dimension framework of trust in AI, similar to Kaplan´s framework (Kaplan et al. 2021), differentiating between human-related, AI-related, and context-related trust. Building upon these previous proposals, our goal is to emphasize a relational view which complements the perspectives that tend to compartmentalize trust. It is important to note that both views are important and irreplaceable, the focus on loci of trust and the focus on relationship between loci. Our proposal tries to acknowledge both, although devoting more time to the relational aspect, which usually is more elusive to analysis. Our underlying purpose is for this clarification work to lead to richer and more precise experimental studies. By analyzing factors that increase, reduce or affect trust in AI, and by empirically measuring, evaluating, and identifying these factors, we believe we can obtain a better understanding of AI-mediated trust in healthcare, and that these insights can inform future empirical research. To this end, we believe that the previous theoretical effort should be translated into a series of impact factors[1] for trust in healthcare AI that can serve as a reference for future empirical work. Therefore, we have distinguished seven specific impact factors for the field of healthcare AI: 1. Function of AI; 2. Type of outcome; 3. Type of pathology; 4. Type of treatment; 5. Healthcare professional's dispositions; 6. Presentation format of the AI; 7. Context of the AI.

For each of these impact factors, we will do the following. After briefly defining each factor, we will explain why and how it impacts trust in the field of healthcare AI. For these explanations, we will rely on the theoretical framework developed in the previous sections, demonstrating how the complex dynamic present between trustor and trustee is concretely reflected in trust in AI, the peculiarities of trust in AI technologies ("AI-mediated trust"), and the eminently relational understanding of trust presented in the last section. While some factors may lean more towards a subjectual or objectual side, we must clarify that none of these factors can properly escape the relational dimension that we have been advocating throughout the article (See Table 1 at the end of this section for a summary of the factors and associated trust constructs).

4.1 Function of AI

The function of AI refers to the specific task that artificial intelligence performs within the healthcare context, such as diagnosis, prognosis, treatment planning, etc. The specific function of AI can influence trust depending on how critical or complex the task is. For example, AI used for diagnosis may generate more distrust than AI used for appointment management due to its direct implications on patient health. Among the previously analyzed constructs, "capacity" (Malle and Ullman 2021) is probably the most important concerning the function of AI, as a fundamental question users will have, at least initially, regarding healthcare

[1] Even if "impact factor" is nowadays associated with "journal impact factor", we consider talking about "impact" is the more intuitive way of referring to the repercussion and significance of certain factors of trust in healthcare contexts.

AI is whether this system can truly perform the task it is being assigned. It is interesting to note that trust in a human doctor is usually integral across all these factors (it would be unusual to fully trust a doctor's diagnostic abilities while completely distrusting their prognoses), whereas with AI, this compartmentalization of trust might occur. Although medical specialization in human doctors contributes to this dissociation of skills and, consequently, of confidence in them, we consider that this dissonance may be much more marked in AI. This factor, like all others, should be fundamentally understood in relational terms but leans more towards the objectual dimension, as it mainly concerns the characteristics of the AI and its performance.

Functions: Diagnosis; Prognosis; Recommendation; Monitoring.

4.2 Characteristics of the output

The characteristics of the output refer to the specific features of the results generated by the AI, which can be precise or imprecise, explainable or unexplainable. The outcomes can also significantly influence the trust generated, whether the results are favorable or adverse, innovative or conventional, etc. This trust-factor is very directly connected with many of the previously analyzed constructs, such as "accuracy" (London 2019), "reliability" (London 2019), or "robustness" (Asan et al. 2020). The impact of this factor on trust is very significant, as it directly affects the relational core of the trust relationship by serving as the basis for potential medical actions to be taken (or not). In this case, the difference between interpersonal trust and trust in technologies could be particularly relevant due to the specificity of problems like interpretability or accountability, which are much more recognizable in interpersonal trust. Similar to the function of AI, this relational factor could be understood as leaning more towards the objectual side since it is more directly related to the characteristics of the results produced by AI. However, it is clear that most of the characteristics attributed to these results can only be so for a subject interpreting them in one way or another, as is especially evident in the cases of adverse or innovative results, to mention one instance.

Characteristics of the output: precise or imprecise; favorable or adverse; innovative or conventional; explainable or unexplainable; interpretable or opaque; accountable or non-accountable; robust or fragile; immediate or long-term.

4.3 Type of pathology

The type of pathology refers to the nature of the disease or medical condition that the AI addresses, such as chronic or rare diseases; but also referring to serious or mild diseases, symptomatic or asymptomatic, etc. One hypothesis is that the accuracy of AI in common pathologies can strengthen trust, while in rare diseases, the lack of data and training cases may generate doubts. It could also be that the severity of a pathology lowers the trust threshold and makes us more likely to try using AI. The construct of "interpretative capacity" (Lipton 2018) can play an important role in a factor like this, precisely because a breakdown in understanding from AI can be perceived, from the human perspective, as a lack of understanding and closeness from the trustee regarding their specific illness. In any case, whatever its specific impact, the factor of the type of pathology, inclined towards the subjectual dimension but no less relational, will probably have a notable influence on trust in healthcare AI.

Types of pathology: serious or mild; rare or common; chronic or acute; asymptomatic or symptomatic; terminal or non-terminal.

4.4 Type of treatment

The type of treatment refers to the kind of medical intervention suggested or administered by the AI, such as medication, surgery, physical therapies, drug administration, etc. Invasive treatments or those involving high risks require a higher level of trust in the AI than non-invasive or preventive treatments. This factor is unequivocally relational, implicating both elements of the trust relationship, trustor and trustee, in an equally unavoidable manner. A factor like this can also be particularly marked by the difference between interpersonal trust and trust in technologies, especially if the degree of intervention in the treatment by the AI goes beyond mere recommendation, such as drug administration by medical robots. Regarding this factor, key constructs like "explainability" (Shin 2021a; Li et al. 2024), "responsibility" (Matthias 2004; Leo and Huh 2020) and "accountability" (Shin and Park 2019) become particularly prominent, as the perception of the proposed treatment's intelligibility and how responsibility for medical decisions is managed will undoubtedly be very important.

Types of treatment: aggressive or conservative; innovative or traditional; surgical intervention or non-surgical intervention; short-term or long-term.

4.5 Healthcare professional's dispositions

The healthcare professional's disposition towards AI includes their willingness, acceptance, knowledge, and previous experience with the technology. The healthcare provider's attitude towards AI can significantly influence the acceptance and use of the technology. Primarily subjectual constructs such as "motivation" or "propensity to trust" would have a specific weight in this trust factor. A positive disposition and deep knowledge of AI improve trust in the system, while ignorance or skepticism regarding AI can create an insurmountable climate of distrust for the use of AI in medicine. Regarding this factor, it also seems clear that constructs like "value alignment" (Israelsen and Ahmed 2019) and "knowledge and previous experience with AI" (Gefen et al., 2003) are of great importance. This issue clearly connects with the specific difficulties of trust in technologies discussed earlier, and although it pertains to the subjectual pole of the trust relationship, it cannot escape relational consideration, as the object (in this case, AI) can inspire trust or distrust depending on its design and performance.

Healthcare professional's dispositions: convinced or skeptical; fearful or calm; attentive or distracted; knowledgeable in AI or not knowledgeable in AI; experienced in healthcare AI or not experienced in healthcare AI.

4.6 Presentation format of AI

The presentation format of AI refers to how AI is presented to the user, including its human or artificial appearance (if the AI is simply presented as a computer application, or if it adopts some human-like appearance such as a humanoid avatar), its robotic or digital form, as well as all possible user interfaces. This factor could also include the distinction between autonomous AIs (something distant, especially in the healthcare field) and those that oper-

ate in conjunction with humans (such as AIs that coordinate with nurses as their advisors). Tangible constructs related to this factor include "anthropomorphism" (Waytz et al. 2014), "humanization" (Cominelli et al. 2021), and "value alignment" (Israelsen and Ahmed 2019) which would likely significantly enhance trust, as well as the interactivity and warmth of the interactions. This is an objectual factor that, once again, undoubtedly also refers to the relational plane, as the appearance of AI matters for how it affects its relationship with human users. Clearly, this factor also involves specific issues of trust in technologies, given that it would represent a divergence from the relatively standardized presentation of human doctors.

Presentation formats of AI: human or artificial; robot or chatbot; text or images; warm or cold; interactive or non-interactive; AI alongside a human or independent AI.

4.7 Context of AI

By AI context, we broadly refer to the characteristics of the environment in which AI is used. This can refer to the physical places where it is employed (medical consultations, nursing homes, private residences, etc.), but also to other contextual elements such as existing laws and regulations, the social and cultural acceptance of AI, etc. Constructs such as "responsibility" (Matthias 2004; Leo and Huh 2020), "value alignment" (Israelsen and Ahmed 2019), and "reliability" (London 2019) prominently appear in relation to this contextual factor. This is an essentially relational factor, as it does not belong to the subjects or objects of the trust relationship but rather to the scenario in which that trust relationship can arise. Although more environmental in nature, this factor has a significant impact on trust in AI. It could be that the use of AI generates more trust in the realm of institutional healthcare management but much less in the clinical field. It could also be that within the clinical setting, there are particularly critical environments, such as surgery, where trust is much more difficult to achieve due to the gravity of the interventions; or fields like nursing, where the human touch could be considered irreplaceable. The underlying issue of regulation and social acceptance could also be decisive, indirectly impacting many of the previously analyzed factors.

AI contexts: hospital or home setting; care or surgical environment; legally regulated or without specific regulation; social acceptance or social rejection.

5 Vignette design for empirical studies

As Vereschak et al. (2021) explain, despite the critical importance of studying the problem of trust in AI, “empirically investigating trust is challenging,” and one of the main reasons is the “lack of standard protocols to design trust experiments” (2021, 1). To conclude this article, we aim to help address this issue by presenting a concrete way in which the preceding sections, with their clarifications and distinctions on trust, could be materialized in an empirical study. Specifically, we propose two sample vignettes, part of an ongoing experimental study on trust, AI and healthcare. While the specific methodological validation of this study remains a work in progress, we believe that the paper’s core contribution is independent of any specific methodology. The theoretical discussion and the trust-factors we describe can be applied across a wide range of contexts, from quantitative to qualitative approaches, and even in training or education-focused studies. The following empiri-

Table 1 Impact factors on AI-mediated trust in healthcare

Impact factors (Types)	Impact factors (Details)	Associated trust constructs
1. Function of the AI	Diagnosis; prognosis; recommendation; monitoring	"Capacity"
2. Type of result	Precise or imprecise; favorable or adverse; innovative or conventional; explainable or unexplainable; interpretable or opaque; accountable or not accountable; robust or fragile; immediate or long-term	"Accuracy"; "Reliability"; "Robustness"
3. Type of pathology	Severe or mild; rare or common; chronic or acute; asymptomatic or symptomatic; terminal or non-terminal	"Interpretative Capacity"
4. Type of treatment	Aggressive or conservative; innovative or traditional; surgical intervention or non-surgical intervention; acute or long-term	"Explainability"; "Responsibility"; "Accountability"
5. Healthcare professional's dispositions	Convinced or skeptical; fearful or calm; attentive or distracted; knowledgeable in AI or not knowledgeable in AI; experienced in healthcare AI or inexperienced in healthcare AI	"Motivation"; "Trust Propensity"; "Value Alignment"; "Previous Experiences with AI"
6. Presentation form of the AI	Human or artificial; robot or chatbot; texts or images; warm or cold; interactive or non-interactive; AI alongside human or independent AI	"Anthropomorphism"; "Humanization"; "Value Alignment"
7. Context of the AI	Hospital or home environment; care setting or surgical setting; legally regulated or without specific regulation; social acceptance or social rejection	"Responsibility"; "Value Alignment"; "Reliability"

cal proposal is just one of many possible ways to operationalize the preceding theoretical framework.

The population sampled in the proposed study will initially be all types of healthcare professionals, even though we plan to expand it to other collectives with the necessary adaptations of the vignettes. To carry out this proposal appropriately, we have tried to follow the warnings and recommendations of Vereschak et al. for conducting empirical studies on trust in AI, avoiding the omission of central elements of trust and distinguishing between different constructs associated with trust (Vereschak et al. 2021, 2), an effort that can be seen in the clarifications and distinctions of constructs and impact factors developed in the previous sections. Specifically, we have followed the recommendation of these authors to investigate the factors of trust between humans and AI, considering the "key elements of trust" (Vereschak et al. 2021, 27). We also follow Reinhardt's suggestion to attempt to measure "possible trade-offs and conflicts between these various values and principles that are supposed to generate trust" (2023, 738), an aspect that, according to this author, has been scarcely addressed in the literature. The vignettes always present complex, multifactorial scenarios that presumably allow for a complex and realistic measurement of trust in AI.

5.1 Sample vignettes

A) Lisa, diagnosed with a severe heart condition, is introduced by her cardiologist to an AI system that recommends a major surgical intervention based on detailed analysis of her genetics, medical history, and symptoms. Although the algorithm suggests this surgery as the best option, the healthcare team shows reservations about using AI. On a scale from 1 to 5, how would you rate your confidence in using this AI system for this situation?

□ 1 (Not at All Trustworthy).
□ 2 (Slightly Trustworthy).
□ 3 (Moderately Trustworthy).
□ 4 (Very Trustworthy).
□ 5 (Extremely Trustworthy).

B) Lisa, diagnosed with a severe heart condition, is introduced by her cardiologist to an AI system that recommends a major surgical intervention based on detailed analysis of her genetics, medical history, and symptoms. The healthcare team, convinced of the accuracy and benefits of AI, fully supports the recommendation and use of AI. On a scale from 1 to 5, how would you rate your confidence in using this AI system for this situation?

□ 1 (Not at All Trustworthy).
□ 2 (Slightly Trustworthy).
□ 3 (Moderately Trustworthy).
□ 4 (Very Trustworthy).
□ 5 (Extremely Trustworthy).

The research question of this vignette could be formulated as follows: "to what extent does explicit support from the medical team influence people's perceived trustworthiness of an AI-driven recommendation for a high-stakes procedure?". Using this vignette, we can show what trust factors were primarily in play, and what factors were secondary to this case. The main factors evaluated in this vignette are "type of pathology"; "type of treatment" and "healthcare professional's dispositions". "Type of pathology" is described as severe and related to a vital organ as the heart; "type of treatment", is described as "surgical intervention", usually considered more serious and intrusive than other minor treatments; finally, "healthcare professional's dispositions" are alternatively described as unexperienced and skeptic in A; experienced and confident in B. Regarding other factors, "function of AI" is not specifically highlighted as it is described generally as technology or tool, encompassing diagnosis and treatment planning. "Characteristics of the output" is also not specified, while the "presentation of AI" and its "context" are left as undetermined. Let us show an additional vignette to give an example of how the trust factors not exposed in our first vignette would be addressed:

A) Charles, a 50-year-old man with type 2 diabetes, began using an AI application recommended by his medical team, led by Dr. Martinez. The application alerted Charles and his doctor to any changes that could indicate deterioration, allowing immediate adjustments to his treatment. Charles felt more secure knowing he had a monitoring device that responded swiftly to his needs. On a scale from 1 to 5, how would you rate your confidence in an AI system's ability to manage monitoring for your medical condition?

□ 1 (Not at All Trustworthy).
□ 2 (Slightly Trustworthy).
□ 3 (Moderately Trustworthy).
□ 4 (Very Trustworthy).
□ 5 (Extremely Trustworthy).

B) Charles, a 50-year-old man with type 2 diabetes, began using an AI application recommended by his medical team, led by Dr. Martinez. The application alerted Charles and his doctor to any changes that could indicate deterioration, allowing immediate adjustments to his treatment. Charles felt more secure knowing he had a monitoring device designed for sustained long-term benefits. On a scale from 1 to 5, how would you rate your confidence in an AI system's ability to manage monitoring for your medical condition?

☐ 1 (Not at All Trustworthy).
☐ 2 (Slightly Trustworthy).
☐ 3 (Moderately Trustworthy).
☐ 4 (Very Trustworthy).
☐ 5 (Extremely Trustworthy).

In this vignette the main factors evaluated in this vignette are "function of AI"; "Characteristics of the output"; "presentation of AI"; "context of AI". "Function of AI" is described as health-monitoring application; regarding "presentation of AI", AI is described as an "application" and as a "device", whereas the "context of AI" is not explicitly mentioned but is clearly implied that the example describes a non-clinical context, but rather a daily context. Finally, "characteristics of the output" are alternatively described as having a real-time, immediate output ("responded swiftly to his needs") in A; and giving "sustained long-term benefits" in B. Regarding other factors, "type of pathology" is explicitly indicated as type 2 diabetes; but "type of treatment" and "healthcare professional's dispositions" are left as undetermined.

This approach can be criticized for mixing several factors, contributing to a possible misidentification of trust problems. However, we respond to this objection in two ways. Firstly, we believe, following Reinhardt's reflection, that we need to measure trust as it really happens in the real world, where we cannot isolate factors and where possible trade-offs between different trust factors are bound to happen (2023, 738). Secondly, the full set of vignettes tries to remediate the possible confusion of trust factors by presenting different scenarios which, through comparison, will allow us to somewhat isolate the strength of each factor.

6 Conclusions

AI in healthcare has the potential to completely transform patient care and clinical decision-making. Trust is known to be essential to technology's effective adoption. Trust in healthcare AI, or as we prefer AI-mediated trust in Healthcare is analyzed in this study. We began with a theoretical approach that made us understand that trust is a complex and relational construct involving not only the technological characteristics of AI systems but also the contextual and individual elements impacting their use. Trust in AI does not merely concern technology's reliability, accuracy, or explainability, but is a complex relational dynamic between the AI, the healthcare professionals, and the patients. This understanding shows the urgency of a comprehensive framework for evaluating, measuring and identifying ways of promoting trust, incorporating subjectual, objectual, and relational dimensions. A framework as the one developed here will help disentangle the complex interplay of factors that

contribute to trust, from the technical robustness and performance of AI systems to the human and contextual elements that shape user perceptions and experiences. We conclude showing how the theoretical discussion and the trust-factors identified could be translated into empirical research. However, it should be emphasized that our specific proposal (the vignettes presented in Sect. 5) represents just one potential application of the theoretical framework, which constitutes the central contribution of this article.

The literature reviewed reveals that trust in AI is profoundly affected by various impact factors such as the function of AI, the characteristics of its outputs, the type of pathology addressed, or the nature of treatments suggested. Special attention should also be given to the dispositions of healthcare providers, the presentation format of AI, and the broader context of its application. These factors collectively influence how trust is formed, maintained, and potentially eroded in medical settings. It is critical to address these issues in order to promote beneficial trust and guarantee that the application of AI technology ends up being truly advantageous.

In the end, developing AI-mediated trust in Healthcare calls for a comprehensive grasp of the socio-technical environment in which these technologies function. We conclude that assessment of the relationship dynamics of trust should be given priority to fully and adequately utilize AI in healthcare while preserving patient rights and welfare. To achieve this goal, we believe that the trust-factor framework exposed in the last part of the article can be instrumental for the purpose of developing more comprehensive and meaningful empirical studies.

Acknowledgements We would like to thank the anonymous reviewers for their thorough revision and for their many helpful and constructive comments that improved the paper significantly.

Author contributions M.A. wrote the main manuscript text.— A.M.A. prepared the introduction and co-wrote the vignette (Sect. 5) with M.A.— R.O.L. prepared the conclusion and co-wrote the table with M.A.— All authors reviewed the manuscript.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. MCIN/AEI/https://doi.org/10.13039/501100011033/; FEDER Una manera de hacer Europa,PID2022- 137953OB-I00,PID2022- 137953OB-I00

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

References

Almyranti M, Sutherland E, Ash N, Eiszele S. (2024). "Artificial Intelligence and the health workforce: Perspectives from medical associations on AI in health", OECD Artificial Intelligence Papers, No. 28, OECD Publishing, Paris, https://doi.org/10.1787/9a31d8af-en.

Asan O, Bayrak AE, Choudhury A (2020) Artificial intelligence and human trust in healthcare: focus on clinicians. J Med Internet Res 22(6):e15154–e15154. https://doi.org/10.2196/15154

Baer MD, Matta FK, Kim JK, Welsh DT, Garud N (2018) It's not you, it's them: social influences on trust propensity and trust dynamics. Pers Psychol 71:423–455. https://doi.org/10.1111/peps.12265

Balkrishnan R, Hall MA, Blackwelder S, Bradley D (2004) Trust in insurers and access to physician: associated enrollee behaviors and changes over time. Health Serv Res 39(4):813–824. https://doi.org/10.1111/j.1475-6773.2004.00259.x

Beck JT, Rammage M, Jackson GP, Preininger AM, Dankwa-Mullan I, Roebuck MC et al (2020) Artificial intelligence tool for optimizing eligibility screening for clinical trials in a large community cancer center. JCO Clin Cancer Inform 4:50–59. https://doi.org/10.1200/CCI.19.00079

Brown HG, Poole MS, Rodgers TL (2004) Interpersonal traits, complementarity, and trust in virtual collaboration. J Manag Inf Syst 20:115–138. https://doi.org/10.1080/07421222.2004.11045785

Calhoun CS, Bobko P, Gallimore JJ, Lyons JB (2019) Linking precursors of interpersonal trust to human-automation trust: an expanded typology and exploratory experiment. J Trust Res 9(1):28–46. https://doi.org/10.1080/21515581.2019.1579730

Caruana R, Lou Y, Gehrke J, Koch P, Sturm M, Elhadad N. (2015). Intelligible models for health care: predicting pneumonia risk and hospital 30-day readmission. Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, https://doi.org/10.1145/2783258.2788613

Cassell EJ (2004) The nature of suffering: and the goals of medicine. 53, 2nd edn. Oxford University Press, Oxford

Castelfranchi C, Falcone R (2010) Trust theory: a socio-cognitive and computational model. Wiley, Hoboken

Choung H, David P, Ross A (2022) Trust in AI and its role in the acceptance of AI technologies. Int J Hum Comput Interact. https://doi.org/10.1080/10447318.2022.2050543

Cominelli L, Feri F, Garofalo R, Giannetti C, Meléndez-Jiménez MA, Greco A et al (2021) Promises and trust in human–robot interaction. Sci Rep 11:9687. https://doi.org/10.1038/s41598-021-88622-9

Davenport T, Kalakota R (2019) The potential for artificial intelligence in healthcare. Future Healthc J 6(2):94–98. https://doi.org/10.7861/futurehosp.6-2-94

Davis FD (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 13(3):319. https://doi.org/10.2307/249008

Ríos Díaz, José, Marcos Alonso, Ramón Ortega Lozano and Aníbal M. Astobiza (Forthcoming). Trust and Acceptability in Medical AI: Psychometric Validation of Clinical Vignettes for Assessing Professional and Public Attitudes Toward Algorithmic Healthcare Interventions.

Floridi L (2019) Establishing the rules for building trustworthy AI. Nat Mach Intell 1:261–263. https://doi.org/10.1038/s42256-019-0055-y

Forghani R, Savadjiev P, Chatterjee A, Muthukrishnan N, Reinhold C, Forghani B (2019) Radiomics and artificial intelligence for biomarker and prediction model development in oncology. Comput Struct Biotechnol J 17:995–1008. https://doi.org/10.1016/j.csbj.2019.07.001

Gefen D (2000) E-commerce: the role of familiarity and trust. Omega 28:725–737. https://doi.org/10.1016/S0305-0483(00)00021-9

Gillath O, Ai T, Branicky MS, Keshmiri S, Davison RB, Spaulding R (2021) Attachment and trust in artificial intelligence. Comput Human Behav. https://doi.org/10.1016/j.chb.2021.106841

Gulshan V, Peng L, Coram M et al (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus Photographs. JAMA 316(22):2402–10. https://doi.org/10.1001/jama.2016.17216

Hall MA, Camacho F, Dugan E, Balkrishnan R (2002) Trust in the medical profession: conceptual and measurement issues. Health Serv Res 37(5):1419–1439. https://doi.org/10.1111/1475-6773.01070

Hatherley JJ (2020) Limits of trust in medical AI. J Med Ethics 46(7):478–481. https://doi.org/10.1136/medethics-2019-105935

Hengstler M, Enkel E, Duelli S (2016) Applied artificial intelligence and trust—the case of autonomous vehicles and medical assistance devices. Technol Forecast Soc Change 105:105–120. https://doi.org/10.1016/j.techfore.2015.12.014

Israelsen BW, Ahmed NR (2019) "Dave… I can assure…you that it's going to be all right…" A definition, case for, and survey of algorithmic assurances in human-autonomy trust relationships. ACM Comput Surv 51(6):1–37. https://doi.org/10.1145/3267338

Jobin A, Ienca M, Vayena E (2019) The global landscape of AI ethics guidelines. Nat Mach Intell 1:389–399. https://doi.org/10.1038/s42256-019-0088-2

Kaplan AD, Kessler TT, Christopher Brill J, Hancock PA (2021) Trust in artificial intelligence: meta-analytic findings. Hum Factors 65(2):337–359. https://doi.org/10.1177/00187208211013988

Kieseberg P, Weippl E, Tjoa AM, Cabitza F, Campagner A, Holzinger A (2023) Controllable AI - an alternative to trustworthiness in complex AI systems? In: Holzinger A, Kieseberg P, Cabitza F, Campagner A, Tjoa AM, Weippl E (eds) Machine learning and knowledge extraction. CD-MAKE 2023. lecture notes in computer science. Springer, Cham

Kim K, Boelling L, Haesler S, Bailenson J, Bruder G, Welch GF (2018) Does a digital assistant need a body? The influence of visual embodiment and social behavior on the perception of intelligent virtual agents in AR. IEEE Int Symp Mixed Augment Reality (ISMAR) 2018:105–114. https://doi.org/10.1109/ISMAR.2018.00039

Kohn SC, de Visser EJ, Wiese E, Lee Y-C, Shaw TH (2021) Measurement of trust in automation: a narrative review and reference guide. Front Psychol 12:604977. https://doi.org/10.3389/fpsyg.2021.604977

Lankton N, McKnight DH, Tripp J (2015) Technology, humanness, and trust: rethinking trust in technology. J Assoc Inf Syst 16(10):880–918. https://doi.org/10.17705/1jais.00411

Leo X, Huh YE (2020) Who gets the blame for service failures? Attribution of responsibility toward robot versus human service providers and service firms. Comput Hum Behav 113:106520. https://doi.org/10.1016/j.chb.2020.106520

Li Y, Wu B, Huang Y, Luan S (2024) Developing trustworthy artificial intelligence: insights from research on interpersonal, human-automation, and human-AI trust. Front Psychol 15:1382693–1382693. https://doi.org/10.3389/fpsyg.2024.1382693

Lipton ZC (2018) The mythos of model Interpretability. Commun ACM 61(10):36–43. https://doi.org/10.1145/3233231

London AJ (2019) Artificial intelligence and black-box medical decisions: accuracy versus explainability. Hastings Cent Rep 49(1):15–21. https://doi.org/10.1002/hast.973

Luhmann, Niklas (1968/1996). *Confianza*. Anthropos: Barcelona.

Malle BF, Ullman D (2021) A multidimensional conception and measure of human-robot trust in trust in human-robot interaction. In: Hancock RH, Billings D, Chen JYC (eds) Cambridge. Academic Press, MA, pp 3–25

Maslej N, Fattorini L, Perrault R, Parli V, Reuel A, Brynjolfsson E, Etchemendy J, Ligett K, Lyons T, Manyika J, Niebles J C, Shoham Y, Wald R, Clark J (2024) The AI index 2024 annual report. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. https://doi.org/10.48550/arXiv.2405.19522

Matheson, R. (2019). Automating artificial intelligence for medical decision-making. MIT News. http://news.mit.edu/2019/automating-ai-medical-decisions-0806

Matthias A (2004) The responsibility gap: ascribing responsibility for the actions of learning automata. Ethics Inf Technol 6(3):175–183. https://doi.org/10.1007/S10676-004-3422-1

Mayer RC, Davis JH, Schoorman FD (1995) An integrative model of organizational trust. Acad Manag Rev 20:709–734. https://doi.org/10.5465/amr.1995.9508080335

Mcknight D, Carter M, Thatcher J, Clay P (2011) Trust in a specific technology: an investigation of its components and measures. ACM Trans Manag Inf Syst 2(2):1–25. https://doi.org/10.1145/1985347.1985353

McLean G, Osei-Frimpong K (2019) Hey Alexa… examine the variables influencing the use of artificial intelligent in-home voice assistants. Comput Human Behav 99:28–37. https://doi.org/10.1016/j.chb.2019.05.009

McLeod, C. (2020). "Trust." The stanford encyclopedia of philosophy. Edward N. Zalta (ed.). https://plato.stanford.edu/entries/trust/

Nittas V, Daniore P, Landers C, Gille F, Amann J et al (2023) Beyond high hopes: a scoping review of the 2019–2021 scientific discourse on machine learning in medical imaging. PLOS Digit Health 2(1):e0000189. https://doi.org/10.1371/journal.pdig.0000189

Nundy S, Montgomery T, Wachter RM (2019) Promoting trust between patients and physicians in the era of artificial intelligence. JAMA. https://doi.org/10.1001/jama.2018.20563

Peete R, Majowski K, Lauer L, Jay A (2019) Artificial intelligence in healthcare. Artifi Intell Machine Learn Bus Non-Eng 2019:96. https://doi.org/10.1201/9780367821654-8

Ramsundar B, Kearnes S, Riley P, et al. (2015). Massively multitask networks for drug discovery. https://doi.org/10.48550/arXiv.1502.02072

Reinhardt K (2023) Trust and trustworthiness in AI ethics. AI Eth 3:735–744. https://doi.org/10.1007/s43681-022-00200-5

Rotter JB (1967) A new scale for the measurement of interpersonal trust. J Pers 35:651–665. https://doi.org/10.1111/j.1467-6494.1967.tb01454.x

Shin D (2021a) Why does explainability matter in news analytic systems? Proposing Explain Anal J J Stud 22(8):1047–1065. https://doi.org/10.1080/1461670X.2021.1916984
Shin D (2021) The perception of humanness in conversational journalism: An algorithmic information-processing perspective. New Media Soc. https://doi.org/10.1177/1461444821993801
Shin D, Park YJ (2019) Role of fairness, accountability, and transparency in algorithmic affordance. Comput Hum Behav 98:277–284. https://doi.org/10.1016/j.chb.2019.04.019
Sim I, Cassel C (2024) The ethics of relational AI — expanding and implementing the Belmont principles. N Engl J Med 391(3):193–196. https://doi.org/10.1056/NEJMp2314771
Sparrow R, Hatherley J (2019) The promise and perils of AI in medicine. Int J Chin Comp Philos Med 17(2):79–109. https://doi.org/10.24112/ijccpm.171678
Thiebes S, Lins S, Sunyaev A (2020) Trustworthy artificial intelligence. Electron Mark. https://doi.org/10.1007/s12525-020-00441-4
Topol EJ (2019) Deep medicine: how artificial intelligence can make healthcare human again. Basic Books, New York
Ueno T, Sawa Y, Kim Y, Urakami J, Oura H, Seaborn K (2022) Trust in human-AI interaction: scoping out models, measures, and methods. In: CHI 2022 - extended abstracts of the 2022 CHI conference on human factors in computing systems, 254. Association for Computing Machinery, New Orleans, LA. https://doi.org/10.1145/3491101.3519772
Venkatesh V, Davis FD (2000) A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 46(2):186–204. https://doi.org/10.1287/mnsc.46.2.186.11926
Vereschak O, Bailly G, Caramiaux B (2021) How to evaluate trust in AI-assisted decision making? A survey of empirical methodologies. CSCW 2021 - The 24th he ACM Conf Comput-Support Coop Work Soc Comput 5(CSCW2):1–39. https://doi.org/10.1145/3476068
Waytz A, Heafner J, Epley N (2014) The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. J Exp Soc Psychol 52:113–117. https://doi.org/10.1016/j.jesp.2014.01.005
Westjohn SA, Magnusson P, Franke GR, Peng Y (2022) Trust propensity across cultures: the role of collectivism. J Int Mark 30:1–17. https://doi.org/10.1177/1069031X211036688
Xing L., Giger M. and Min J. (2021). Artificial intelligence in medicine. Technical basis and clinical applications. Academic Press. London
Zuboff S (1988) In the age of the smart machine: the future of work and power. Basic Books, New York