



Adopting trust as an ex post approach to privacy

Haleh Asgarinia¹

Received: 15 July 2023 / Accepted: 7 January 2024
© The Author(s) 2024

Abstract

This research explores how a person with whom information has been shared and, importantly, an artificial intelligence (AI) system used to deduce information from the shared data contribute to making the disclosure context private. The study posits that private contexts are constituted by the interactions of individuals in the social context of intersubjectivity based on trust. Hence, to make the context private, the person who is the trustee (i.e., with whom information has been shared) must fulfil trust norms. According to the commitment account of trustworthiness, a person is trustworthy only if they satisfy the norm of competence. It is argued that a person using an AI system to answer a question is competent only if they are ex post justified in believing what has been delivered by the AI system. A person's belief is justified in the doxastic sense only if the AI system is accurate. This feature of AI's performance affects a person's competence and, as a result, trustworthiness. The effect of AI on trust as an essential component of making the context private, and thus on privacy, means an AI system also impacts privacy. Therefore, a private context is constituted when the individual with whom the information is shared fulfils the competence norm and the AI system used for analysing the information is sufficiently accurate to adhere to this norm. The result of this research emphasises the significance of the relationship between individuals involved in information-sharing and how an AI system used for analysing that information impacts the relationship regarding making the context private, as well as how it impacts privacy. The findings of this research have significant implications for improving or ameliorating privacy regulations in light of trust.

Keywords AI accuracy · Competence · Social value of privacy · Trust norms · Trustworthiness

1 Introduction

Privacy can be understood as a social construction we create as we negotiate our relationships with others on a daily basis. By placing privacy in the social context of intersubjectivity [1], privacy is conceived as a dynamic process regulating interpersonal boundaries by drawing a negotiated line between openness and closedness to others [2, 3]. The dialectical approach to privacy, in which privacy can only be obtained through the negotiated interaction between social actors, captures its importance as a social value [1, 3].

The dialectical approach to privacy neglects that privacy is a social phenomenon not only because other people exist, but also because privacy concerns the social circumstances in which information flows from one party to another. The

contextual integrity model of Nissenbaum [4] elaborates on socially embedded privacy in the digital age. Nissenbaum argues that different social contexts are governed by different social norms that govern the flow of information within and outside of that context. Protecting privacy entails ensuring the appropriate flow of information between and among contexts. Privacy is a norm that regulates and structures social life [4].

According to Waldman [5], although Nissenbaum [4] succeeds in the socialising theory of privacy in terms of social interactions and the possibility for individuals to be properly embedded in social relationships, it begs the question of what a 'private context' is. Waldman responds to this question by arguing that 'private contexts are defined by relationships of trust among individuals' [5, p. 559].

Drawing on the insights of Waldman [5], a private context is constituted by relationships of trust among the individuals involved in the context. The interaction of different individuals in the social contexts of intersubjectivity based on trust constitutes privacy. Privacy is a social construction

✉ Haleh Asgarinia
h.asgarinia@utwente.nl

¹ Department of Philosophy, University of Twente,
Drienerlolaan 5, 7522DB Enschede, Netherlands

we cannot have unless we work together, which is what Altman [3] and Steeves [1] argue. Interpersonal trust depends upon the nature of relationships between individuals, social circumstances, and context. Since privacy depends on trust, such social circumstances are associated with the value of privacy as well, as Nissenbaum [4] considers in her socialising theory of privacy.

Ex post approaches discuss privacy when information is shared and revealed between different individuals in a context. Trust as an *ex post* approach to privacy, as highlighted by Waldman [5, 6], emphasises the role of individuals in constituting a private context. Accordingly, privacy scholars have been working on trust norms and have regulated trust-promoting norms that govern the relational duties of trustee parties (i.e., the person who is trusted) regarding how to build and cultivate trust-based relationships with trustors (i.e., the one who trusts), thereby making the context suitable for disclosures. Richards and Hartzog [7], for example, have identified trust norms, such as protection, discretion, honesty, and loyalty.

This paper adopts a philosophical perspective to identify trust norms, which differ from those in, for example, the work of Richards and Hartzog [7]; such studies often consider trust and privacy from a legal perspective. From a philosophical perspective, as this paper argues, competence must be considered a norm to be trustworthy. Thus, the norm of competence must be included in the list of trust norms that Richards and Hartzog [7] have proposed.¹ Moreover, this paper emphasises the role of AI systems in establishing a person's trustworthiness and in contributing to making the context private. This paper explores the significance of AI in contributing to B's trustworthiness and thereby constituting private contexts, a topic that has not been given adequate attention in the literature.

For a clearer understanding of cases in which both individuals and an AI system are involved and information is shared and revealed, consider the following case:

A person (B) uses data (q) about another person (A) to predict whether she has breast cancer (p). B cannot deduce if q then p ($q \rightarrow p$) because of his limited background knowledge. To deduce if q then p, B relies on a machine learning (ML) model to identify the possible presence of breast cancer for A. Such an ML model has displayed the potential to predict whether A develops breast cancer within certain timeframes by analysing her electronic health records and mammography patterns [8]. The deliverance of the ML model is a propo-

sition in response to the following question: 'Is breast cancer present?'

Trust as an *ex post* approach to privacy emphasises the role of B, as a person trusted by A, in constituting a private context. In addition to B acting as the trustee, does the ML model, which is used to predict whether p, contribute to making the context private? This paper argues that, yes, ML models that predict aspects such as the presence of breast cancer contribute to making the context private. Furthermore, the ML model impacts trust relationships between A and B. Since privacy depends on trust, the ML model contributes to making the context private, ultimately impacting privacy considerations. Therefore, adopting trust as an *ex post* approach to privacy not only emphasises the role of the trustee, as Richards and Hartzog [7] highlight, but also the role of AI systems in constituting private contexts.

The main purpose of this research is to investigate how an ML model affects privacy. Since AI systems based on ML develop ML models, the main research question (RQ) is as follows: 'How does an AI system affect privacy?' To respond to this, I formulated two sub-questions (SQs): 1. 'How do A and B cultivate or maintain relationships of trust?' 2. 'How does an AI system affect trust relationships between A and B?' Answering these two SQs provides the foundation for answering the main RQ. The SQs are addressed in Sects. 2–3, respectively. Each response forms a premise for the argument that concludes with an analysis of the impacts of AI on privacy. The assumptions and premises of the argument that I formulate in this paper are presented below.

I assume that privacy is constituted by the interaction of different individuals in the social context of intersubjectivity based on trust. Privacy, in a disclosure context in which information is shared and revealed, can thus metaphorically be conceived as a realm constituted by trust-based relationships. Hence, cultivating trust in a context is essential to making that context private. Additionally, I focus on cases in which A shares data (q) with B to answer a specific question, and B responds to the question based on the AI-delivered proposition p. As a result, the particular task that B is relied upon to perform is to assert p.

Section 2 addresses SQ1, which establishes the first premise of the argument. It is argued that, to promote trust in a context, A and B must conform to trust norms. B can be trustworthy while avoiding unfulfilled commitments. Given that promise-making norms are the most explicit mechanism by which B takes a (new) commitment, norms of being trustworthy derive from norms regarding promise-making. Competence is one of the norms of promise-making [9]. As a result, trusting B's words involves relying upon him to fulfil promise-making norms, including the norm of competence. Section 3 addresses SQ2, which forms the second premise

¹ Unlike Richards and Hartzog [7], who consider privacy solely as secrecy and formulate trust norms based on this definition, I conceive of privacy as a social phenomenon constituted by trust, and I formulate trust norms beyond merely secrecy.

of the argument. It is argued that an AI system affects B's competence and, thus, the trust relationships between A and B. Finally, given that AI affects trust, which is a constituent component of privacy, Sect. 4 concludes that AI affects privacy. Since trust requires an accurate AI system, privacy does also.

2 Trust

How do A and B cultivate or maintain relationships of trust? To promote or preserve trust in the context, A and B must conform to trust norms. To identify trust norms, I consider interpersonal trust rather than trust in a group or institutional trust, and I adopt four assumptions.

First, trust is a three-place relationship involving two people and a task. According to the majority of the literature [9–14], trust is generally a three-place relation: A trusts B to φ . A primarily trusts B to do some particular thing rather than trusting him in general and in every way. Second, I focus on the norms of trust from the trustee's side. Norms of trust arise between two parties: a norm to be trusting in response to the invitation to trust and a norm to be trustworthy in response to the other's trusting reliance [15]. The former norm lies on the trustor's side, and the latter on the trustee's side [16]. In this paper, I discuss the norms of trust on the trustee's side and the conditions that give rise to trustworthiness in three-place relations. Third, I adopt doxastic conditions on trust. According to doxastic accounts, trust involves a belief on the part of the trustor. When A trusts B to φ , A *believes* that B will φ [13]. Fourth, like most philosophers, I distinguish trust from mere reliance. Trust involves reliance 'plus some extra factor'. Controversy surrounds this factor, which generally concerns why the trustor would rely on the trustee to be willing to do what they are trusted to do [12, p. 5].

Regarding the first assumption, trust can be a two-place or a three-place relationship. It is a relationship between a trustor and a trustee in the first instance, as in A trusting B. Two-place trust, as opposed to three-place trust, is fundamental, according to Faulkner [17]. Two-place trust is a rather demanding affair; when we state that A trusts B *simpliciter*, we ascribe A a rather robust attitude, one in which A trusts B in several respects. A three-place relationship, on the other hand, is a less-involved affair: when we state that A trusts B to do φ , or that A trusts B with a valued item C [10], we do not need to express much about their relationship. According to Carter and Simion [16] views in 'The Ethics and Epistemology of Trust', this difference is maintained when we focus on the trustee's trustworthiness. One can be trustworthy in general, but one can also be trustworthy regarding a particular matter. I think of a trust-based relationship as a three-place relation between two people and

a task. Considering three-place trust to be a general relation of trust, B can be trustworthy with regard to a particular matter but not generally. For example, B can be trustworthy in keeping a meeting appointment but may not be trustworthy overall. With respect to the case discussed in this paper, B can be trustworthy with regard to the task of assertion.

According to the second assumption, I only consider the norms of trust on the trustee's side. Addressing the third assumption, in discussions regarding the rationality of trust, or whether the trust is appropriate or well-founded, it is crucial to explore whether trust essentially involves belief. Proponents of non-doxastic accounts, such as Holton [14], argue that it is not essential for trust to involve a belief about the trustee, such as a belief that they are trustworthy. Jones [18], for example, maintains that the trustor must have an affective attitude which is not described by belief. Trust involves affective attitudes that may lead to corresponding beliefs. Hence, the rationality governing trusting is distinct from rational belief. However, proponents of doxastic accounts, such as Hieronymi [13] and Hawley [9], argue that trust involves a belief on the part of the trustor. Hence, if trust is a belief, the rationality that governs trusting is drawn from rational belief. To the extent that the trustor is rationally entitled to believe that the trustee is trustworthy with respect to φ , the trustor thereby has an entitlement to trust the trustee with respect to φ [16]. I defend a doxastic account of trust mainly because it requires less explanation as to why trusting someone would give us a reason to believe what they say; 'trust gives a reason for belief because belief can provide a reason for belief' [19, p. 113]. Although discussions of the entitlement to trust and the rationality of trust are important, I do not address them because these subjects are more related to trust norms on the trustor's side than on the trustee's side. I simply assume that, for trust to be well-grounded, the trustee must be trustworthy.

Is it required for A to have evidence of B's trustworthiness to be entitled to trust B? According to Hinchman, A's trust in B is reasonable even if A has no evidence of B's trustworthiness on the relevant matter, but it is not reasonable if A has good evidence of B's untrustworthiness on that matter. It is in line with the externalist approach to trust that the trustor need not have access to or be aware of the evidence [20, p. 580]. I agree with Hinchman's [20] point that reasonable trust does not require evidence of B's trustworthiness to be available to the trustor. Again, while the rationality of trust is important, most discussion on it is focused on the trustor's side.

Finally, concerning the fourth assumption, Baier [10] provides an influential account of trust. According to her, trust must be distinguished from mere reliance. Although we can rely on both people and inanimate objects, not everything can be genuinely trusted. Trust differs from mere reliance because, when an object breaks, one may be disappointed,

but one does not feel betrayed. However, when we trust and are let down, we feel betrayed. As expressed in Hieronymi's [13] theory, trust requires something more than merely relying on someone to do something; it requires a vulnerability to betrayal if let down.

Most philosophical theories of trust [9–12, 14] are explicitly designed to explain that trust is a form of reliance, but it is not mere reliance; rather, trust involves reliance 'plus some extra factor' [12, p. 5]. Different theories associate this extra factor with the motives of the trustee. If A trusts B to φ , then A relies upon B to φ ; moreover, A assumes B has the right motive for φ -ing [10, 11]. Those theories that dispute what type of motive the trustee should have to make trust appropriate are classified as 'motive-based' theories [12]. The other category of theories associates the extra factor with the trustor's particular stance towards the trustee [9, 12, 14]. These theories are classified as 'non-motive-based' theories, according to McLeod [21]. In what follows, I explore whether motive-based or non-motive-based theories succeed in explaining the conditions that give rise to trustworthiness.

I begin my argument by considering the task that B is relied upon to perform in general as φ . In sub-Sect. 2.2.2, I specify φ . Regarding this, those who are concerned with the task of assertion might skip the first three sub-sections and move to the last one.

2.1 Motive-based theories on trust

According to motive-based theories, the conditions that lead to trustworthiness are based on the motivation a trustworthy person has. Goodwill or self-interest are two examples of such motivations.

A trustworthy person is motivated to act by virtue of their goodwill towards the trustor. According to Baier [10], when we trust someone, we rely on them having goodwill towards us. However, Holton [14] argues that Baier's goodwill account of trustworthiness is not absolutely correct. Primarily, relying on a person's goodwill towards oneself is not a sufficient condition for trust. A confidence trickster might rely on your goodwill without trusting you. Second, goodwill is not a necessary condition: I can trust a person without relying on their goodwill towards me. I can, for instance, trust someone to look after a third party without requiring them to have goodwill towards me.

Another motive-based theory describes trustworthy people's motives in terms of self-interest, such as in the encapsulated interests account of Hardin [11]. He contends people trust those they believe have strong reasons to act in our best interests. He claims the primary motivation of individuals we trust is to preserve a relationship with us. Trustworthy people are motivated by their own interest in maintaining the relationship they have with the trustor, which motivates them to encapsulate that person's interests in their own.

McLeod [21], however, provides an example to demonstrate why Hardin's [11] theory is flawed. Consider a sexist employer who is interested in maintaining relationships with female employees and treats them fairly but whose interest derives from a desire to keep them around to daydream about having sex with them. This interest conflicts with the women's interest not to be objectified by their employers. At the same time, if the women were unaware of his objectification of them, he could ignore this particular interest of theirs. He can maintain his relationships with them while ignoring their interest in not being objectified, and encapsulating enough of their other interests in maintaining a good relationship in his own. This situation, according to Hardin, would make him trustworthy. However, if the women knew the main reason for their employment, they would not find him trustworthy. Being motivated by an interest to maintain a relationship may not require adopting all the trustor's interests to be considered trustworthy by that person.

Although motive-based theories are not limited to goodwill and self-interest theories, these are the dominant viewpoints in the literature. However, since these theories do not provide an appropriate account of trustworthiness, we need other theories that identify conditions for being trustworthy that are not driven by goodwill or self-interest.

2.2 Non-motive-based theories on trust

The conditions that lead to trustworthiness reside in the stance the trustor takes towards the trustee [21]. One can be trustworthy while avoiding unfulfilled commitments, regardless of one's motivation for fulfilling commitments. A relies on B to φ because A believes B has a commitment to φ -ing [12].

Holton [14], like Baier [10], distinguishes between trust and mere reliance. However, unlike Baier, he does not suggest that, when we trust someone, we rely on them to have goodwill towards us; instead, when we trust someone, we take a particular stance towards them, which is the participant stance. Holton highlights that, in addition to resentment and gratitude, the feeling of betrayal is one of what Strawson [22] calls the reactive attitudes. We normally take these attitudes towards people but not towards objects. Behind these classes of attitudes is a more general attitude, which Strawson calls the participant attitude and Holton calls the participant stance. The participant stance is a particular reactive attitude we take towards those we regard as responsible agents. When we interact with someone who provokes a reactive attitude, whether resentment or gratitude, we adopt a particular attitude that is bound with the ascription of responsibility towards them. According to Holton [14], trust is a reliance on the participant stance: trust involves something like a participant stance towards the trustee. Despite Holton's [14] correct identification of the participant stance

as a required component of trust, Hawley [12] finds Holton's theory unsatisfying because relying upon someone to whom you take a participant stance does not always entail trusting them; some interactions occur outside the realm of trust.

According to Hawley's [12] view, which she elaborates on in her book *How to be Trustworthy* [9], it is reasonable to trust someone to do something only if that person has an explicit or implicit commitment to doing it. To trust someone to do something is to believe they have a commitment to doing it, and to rely upon them to meet that commitment. To make her account plausible, Hawley employs a very broad notion of commitment. Commitments can be implicit or explicit, weighty or trivial, conferred by roles and external circumstances, default or acquired, welcome or unwelcome. Hawley's account of trustworthiness in the context of the commitment, in terms of avoiding unfulfilled commitment, has nothing to do with the trustee's motives. To be trustworthy in some specific respect, it is enough to behave in accordance with one's commitment, regardless of motive. One person may trust another to do something without believing them to be motivated by their commitment [12, pp. 10–11, 16]. In what follows, I adopt the commitment account of trustworthiness and identify norms to be trustworthy in response to the other's trusting reliance.

2.2.1 The commitment account of trustworthiness

According to Hawley [9, 12], commitment is at the centre of the notion of trust. The most explicit mechanism through which we take on (new) commitments is promise-making. When thinking about promises and trust, two questions arise: first, how do we decide whom to trust? Second, whose promises do we accept and rely upon? The first question is from the perspective of the promise-receiver, whereas the second is that of the promise-giver. The following argument focuses on the second perspective and answers the following question: 'What do good promisors do?' In Hawley's [9] view, good promisors not only keep their promises, but they also make appropriate promises in the first place. Making a good promise requires a sincere intention, the permissibility of the action promised, and the competency to keep the promise. Hence, the norms regarding promise-making are sincerity, promising to act morally, and competence. Among these norms, I focus on competence as it is impacted by AI, a topic that is discussed in Sect. 3.

A good promise requires competence to keep the promise, which is a norm of promise-making: do not make promises you are not competent to keep. 'Competences are dispositions of an agent to perform well', and they have three components: constitution, condition, and situation [23, p. 465]. Similarly, the competence required to keep a promise includes these three components [9]. After

explaining competence, I return to the competence norms for promise-making.

Consider colour vision competence in Sosa's [23] paper on 'How Competence Matters in Epistemology', for instance. A constitution competence includes rods and cones; a condition competence includes being awake and sober; and a situation competence includes adequate light. When a person's visual systems are fully functional, they are awake and alert and they see the object in plain view, exercising colour vision competence. Not only does a person need competence in colour vision, they also need the competence to assess the required conditions and situation of the proposed competence—second-order assessment. According to Sosa [23], then, an agent's success relies not only on their constitutional competence, but also on their being in an appropriate shape while appropriately situated. Thus, an internal constitution, being in good shape to exercise that competence, and external circumstances are required if the performer is to be properly credited with complete competence.

Analogously, the competence required for good promise-making should encompass all three components. In Hawley's [9] view, constitutional competences include a steady, reliable capacity to achieve success. More precisely, I argue the notion of competence is close to exercising a reliable intellectual capacity to form a justified belief. That is, keeping the promise manifests competence when forming an epistemically justified belief that one will keep the promise to ϕ . The following paragraphs further clarify what the constitution competence of promise-making involves.

The second component of competence, the condition competence, requires a person to be awake, alert, and sober when making a promise. The third component of competence, the situation competence, indicates that what we are competent to do depends on the external circumstances, including the physical environment, social environment, and material resources we find ourselves in. Therefore, to incur a certain commitment, we require insight not only into our capability or underlying skills, but also into an actionable feature of our environment. For instance, it is far more difficult for a doctor working in a field hospital than it is for someone working in a well-equipped hospital to save a child's life. In a challenging environment, the situation competence required for success differs from in an easy environment. Acting in different environments requires a doctor to use different competences, some of which are more difficult to develop and maintain than other. Therefore, a person who makes a commitment needs to be aware of the circumstances in which they will need to act [9].

I have described how a person being competent to promise to ϕ depends on being in good shape while making a promise and the complex facts regarding their physical and social environment. Now, I return to the first component

of competence: constitutional competence. To possess a corresponding competence to keep a promise, I argue one should be ‘justified’ in thinking they will keep a promise when making one. Goldman [24] distinguishes two uses of ‘justified’: an ex post use and an ex ante use. The ex post use occurs when there exists a belief, and we say whether that belief is justified. The ex post or doxastic sense of justifiedness applies to beliefs that a subject actually holds, rather than beliefs they could hold. In contrast, the ex ante use occurs when no such belief exists, or when we wish to ignore the question of whether such a belief exists. The ex ante or propositional justification applies to a proposition (p), a subject, and their epistemic situation. If we say that a subject is propositionally justified regarding p, we mean that it would be appropriate for them to believe p; it is applicable even if they have no belief in the specified proposition [25, p. 5]. Since I argue it is inappropriate to promise to ϕ while one does not possess a belief in ϕ , I use an ex post or doxastic sense of justifiedness. Therefore, I articulate a good promise regarding satisfying the competence norm is one in which the promisor *in fact* believes they will keep the promise, rather than believes it is *possible* to keep the promise, and their belief is *justified*.²

In the scholarly literature, there are different theories of doxastic justification, such as *mentalist evidentialism* and *process reliabilism*. Since I adopt the externalism approach in this paper,³ I focus on the process reliabilism theory of justification. The justificational status (J-status) of a belief, according to this theory, depends on how it is formed, or caused. As the theory indicates, how a belief is causally produced is crucial to its J-status [26]. Consequently, the competence required for making a promise includes a capacity to form a belief based on a reliable process. The reliabilist principle of justification can be explained as follows:

(R) A belief B (at time t) is justified if and only if B (at t) is the output of a series of belief-forming or belief-retaining processes, each of which is either unconditionally or conditionally reliable, and where the conditionally reliable processes in the series are applied to outputs of previous members of the series. [26, p. 35]

I argue that competence includes a reliable capacity to form a justified belief to achieve success in what is promised. When a person promises to ϕ when they lack such a capacity, they make a wrong promise. Consider the following

example provided by Hawley [9]. A child, Cindy, is brought to the hospital with a sever and an unfamiliar condition. The junior doctor in charge of the case, Jack, promises the parents he will save their child’s life, and he sincerely intends to do so. Cindy’s condition can be treated with a certain type of antibiotic, which Jack happens to try first, saving Cindy’s life. In this case, the junior doctor genuinely intends to save Cindy’s life, which is morally permissible. Jack keeps his promise but only through sheer luck, rather than through his competence. He does not have a justified belief he would keep the promise. Therefore, his promise counts as over-promising. For simplicity, I presume the doctor was awake, and I do not consider whether he was at risk of lacking the situation competence. I only concentrate on the requirement not explored in depth in Hawley’s [9] description, which is having a justifiable belief in accomplishing the promised action or activity.

Even if Jack believed he would save Cindy’s life, his belief, in the doxastic sense of justifiedness, would not be justified. Although Jack has no outstanding skills regarding diagnosing and treating such conditions, he promised he would save Cindy. This promise was merely wishful thinking, but it made him confident. According to Goldman [26], wishful thinking is a highly flawed thought process. Forming belief through wishful thinking is unjustified, meaning Jack’s belief was unjustified. Since competence includes being justified in believing what is promised, Jack was incompetent in this case. However, as Hawley [9] points out, a lack of suitable competency does not imply incompetence in the normal sense. Jack was as competent a doctor as his peers, but he was not competent to save Cindy’s life in this circumstance.

Consider another case identical to the previous one, except Jill, the senior doctor, is substituted for Jack. Jill is an experienced physician and promises the parents she will save Cindy’s life, which is what she sincerely intends to do. She has an idea about the condition the child is suffering from, and whether it is treatable. In this case, Jill arrived at the justified belief she will save Cindy’s life (B) by drawing inferences from her old belief. She acquired this belief from reading a medical journal (M) that reported a patient with Cindy’s symptoms was treated in a specific way (x). Jill also believed M is very trustworthy in such matters, based on her experience. Jill’s belief in curing specific diseases was stored in her memory and accessible to her. She made an inference from the belief retained in her memory and believed she would save Cindy’s life.⁴

Following (R), Jill’s belief in saving Cindy’s life is justified because it is an output of a reliable process (inferential process) involving reliable inputs. Jill first used perceptual

² In my view, a person is rationally permitted to perform ϕ only if the person has a justified belief in ϕ .

³ As I mentioned in Sect. 2, I suppose it is not required for A to access the reasons contributing to B’s trustworthiness. Since I take an externalist reading of reasons to believe B’s trustworthiness on the trustor’s side, I do the same when analysing competence norms on the trustee’s side.

⁴ One of Goldman’s [26] examples inspired me to make such a case.

processes to form the belief that M reports the specific cure. The perceptual step is unconditionally reliable. According to (R), a belief is justified if it is produced by a belief-forming process that is unconditionally reliable. Jill then inferred from experience that M is trustworthy enough for her belief to be true regarding the specific disease cure. The inferential step is conditionally reliable. According to (R), the belief is produced by the inference process, which is a conditionally reliable process, and the input of this process, that is, her old beliefs, is justified. Next, the memory stage is a conditionally reliable belief-retaining process; its later outputs are usually true if the earlier inputs to it were true. Finally, she used the inferential step to infer that she would save Cindy's life. As I mentioned previously, the inferential step is conditionally reliable. Then, using principle (R), Jill's preserved belief in B is justified. Promising that she will save Cindy's life is a good promise as it meets the (internal) requirements of the competency to keep the promise. In contrast to Jack, Jill is competent to make the promise she will save Cindy's life.

In summary, A trusts B to φ because A believes B has a commitment to φ -ing. To be trusted when making a commitment, B must comply with the norm of promise-making. A good promise requires competence. The constitutional competence required for making a commitment is that B is *ex post* justified in believing he will successfully φ . Next, I specify the task (φ) that B is relied upon to perform as an assertion, and I analyse the norms associated with this specific task.

2.2.2 Assertion as promising: the norms of being a trustworthy assessor

I have explained that, to promote trust, a person who is trusted to perform φ must conform to the norm of competence. Since the main case in this paper is the one in which the task that the trustee is relied upon to perform is to assert p — φ is specified with an assertion—this section explores the norms of being a trustworthy assessor.

What norms must a person meet to be a trusted assessor? Trusting other people's words involves relying upon them to fulfil a commitment, to satisfy both *promise-making* and *promise-keeping* norms. A commitment made by a speaker when making an assertion is that they speak justifiably. When *promise-making* norms are applied in the context of assertion, trustworthiness is required as competence in speaking justifiably. When the *promise-keeping* norm is applied to the assertion, the trustworthy assessor must in fact speak justifiably. This section clarifies the norms of being a trustworthy assessor [9].

Asserting or telling⁵ involves a form of promise. One way to think of assertion as a special case of promising is to identify asserting that p to promising to p . In other words, since asserting involves a form of promise, it is a promise to p . Therefore, asserting that p is identical to promising to p . For example, when someone asserts there is snow outside, they promise there is snow outside. However, Hawley [9] maintains it is unacceptable to identify asserting p to promising to p . When making an assertion, one need not be in a strong epistemic situation, such as when making a promise. By asserting that p , one does not become obliged to make it true that p . Thus, the account of assertion regarding promising does not entail identifying an assertion p with a promise to p .

Hawley [9] proposes another way to assimilate assertion to promise by working out what a person is promising to do when making an assertion. She claims asserting whether p involves both

- (a) promising to speak truthfully regarding whether p ; and
- (b) speaking truthfully or untruthfully regarding whether p (i.e., keeping or breaking the promise).

Before proceeding, I modify Hawley's account of assertion regarding the promise. The idea that identifies assertion to promise emphasises that assertion entails making a claim about something in fact in the world. This idea is rejected by Hawley [9], who instead defends the idea that assertion involves a promise to speak in ways that match the world; a promise to speak truthfully requires promising there is a match between words and the world [9, p. 52]. Truth, in both propositions, that is, either there is something in the world or that words are matched to the world, is a purely metaphysical concept rather than an epistemological one. In both claims, what makes the proposition true or false is simply the state of the world. The claim's truth value is not affected by the cognitive relations people have towards the relevant state of affairs. However, I state that assertion involves a promise to speak *justifiably*, which requires promising there is a cognitive relationship with the relevant state of affairs asserted. As Goldman highlights, 'cognitive relations to a proposition are crucial for determining justification or warrant. A person's justifiedness with respect to speaking as to whether p is never (or rarely) fixed by its actual truth value' [25, p. 5]. Given the difference between taking a claim to be justified and taking it to be true, I believe assertions are not faulty if the speaker lacks any evidence for its truth; rather, it is

⁵ I continue to use the term 'assertion' rather than 'telling' because it is consistent with the terminology employed by Hawley [9] and Brandom [27], whose works serve as the foundation for this section.

possible to have highly favourable evidence that justifies a proposition despite its falsity.

To address the concern related to the notion of truth in Hawley's account, I propose the following requirements: asserting regarding whether *p* involves both

- (a) promising to speak justifiably regarding whether *p*; and
- (b) speaking justifiably or unjustifiably regarding whether *p* (i.e., keeping or breaking the promise).

There are three points to note about treating assertion as promising. First, Hawley's [9] view differs from a Brandom-style commitment [27] to justify *p*. Second, as the condition of b) in Hawley's account and the corresponding condition in my view illustrate, making a promise and keeping or breaking it happen simultaneously in the case of assertion. Third, the norms of promise, including competence is applied in the case of assertion. I now discuss each of these points in detail.

First, Hawley's account of assertion in terms of promising to speak truthfully (or justifiably, in my view) differs from a Brandom-style commitment. According to Brandom [27], asserting a sentence entails a commitment to present a justificatory defence of it. Brandom suggests that 'the commitment involved in asserting is to undertake the justificatory responsibility for what is claimed. In asserting a sentence, one commits oneself to justify the claim' [27, p. 641]. Assertions are treated as warranted until challenged. One commits oneself to justify assertions once a specific question is raised regarding them. Although there is no end to the justification of the justification, and each justifying assertion may be questioned and need additional justifying assertions, the assertor must provide an appropriate set of justifying assertions if challenged [27]. For example, I assert there is snow outside to my neighbour. In responding to a challenge by my neighbour that the white stuff is not snow, but foam, I assert I saw no person, or film crew, put foam outside. Hence, I provide a set of justifying assertion(s) inferentially related to the original claim.

However, Hawley [9] contends that assertion does not involve commitments that extend beyond the moment of making the assertion, either in terms of justification or retraction. In Hawley's view, people who make a promise to do something become obliged to do it, but they do not become obliged to provide evidence of having done so if challenged. For example, if a son promises his mother he will finish his homework before dinner, he is obliged to do so. Nevertheless, he is not obliged to show his mother the completed homework. The son refuses to show his schoolwork because he wants his mother to trust him, to take him at his word. Otherwise, his mother's inability to relax reveals a lack of trust. Trusting someone to keep their promises typically involves relying upon them to behave in the manner in which they committed to behaving and does not involve

justificatory commitments. Similarly, a promise to speak truthfully (or justifiably, in my view) does not require an assertor to provide evidence they have spoken truthfully (or justifiably, in my view) even if challenged [9]. I agree with Hawley in that I think an account of assertion in terms of promise does not entail anything as extensive as Brandom's commitment account of assertion.

Second, assertion involves a promise to speak justifiably⁶ and keeping or breaking that promise at the same time. The promise made in assertion is uncommon because it is made and kept at the same time, or else made and broken at the same time. For example, Clara asks Emma, 'Do you promise to say your next word as loudly as you can?' Emma shouts back, 'YES!' Emma promises to speak as loudly as she is able, and then simultaneously either keeps or breaks the promise. The promise to speak justifiably is kept by speaking justifiably [9]. An assertor keeps the promise to speak justifiably once they are speaking.

Third, as I mentioned in Sect. 2.2.1, promise-making is governed by the norm of competence. When the norm is applied to the special case of promising to speak justifiably regarding whether *p*, the following result is obtained:

- (a) One must promise to speak justifiably regarding whether *p*:
 - only if one is competent to speak justifiably regarding whether *p*.

What elements are required for a competence norm in a promise to speak justifiably regarding whether *p*? Remembering that proper promise-making requires competence, the response in this respect is to not promise to speak justifiably regarding whether *p* unless you are competent to speak justifiably regarding whether *p*. Competences, according to Sosa [23], encompass three components: constitution, condition, and situation. I begin with the latter two components and return to the first afterward. In the case of assertion, conditional competence is achieved when the assertor is sober, awake, and alert. Situational competence is related to the circumstances in which an assertor must act or speak justifiably. Regarding the specific task of assertors to utter *p*, the external circumstances might be to ensure what audiences expect to hear from them, as indicated by Hawley [9].

Constitutional competence, I argue, is close to the doxastic sense of justification for what is promised. More precisely, I claim that one is competent to keep a promise to ϕ when one is doxastically justified in believing ϕ .⁷ Similarly,

⁶ Although the term 'truthfully' is used in Hawley's argument, I use the term 'justifiably' in the remainder of the paper for the reasons stated above.

⁷ In this paper, I am not discussing whether testimony transmits knowledge (or justification) or generates knowledge (or justification).

one is competent to speak justifiably regarding whether p, only if one is ex post justified in believing whether p. Consequently, one has the appropriate competence to assert whether p only if one is ex post justified in believing whether p.

My view is that one is competent to assert whether p only if one justifiably believes whether p, which differs from the *reasonable to believe norms of assertion* proposed by Lackey [28]. Lackey highlights that one should assert whether p only if it is reasonable for one to believe whether p. According to Lackey, an assertor might fail to believe whether p; nevertheless, they have substantial evidence indicating that such a proposition should be believed, rendering it reasonable for them to believe whether p [28, p. 125]. However, I claim, to be competent in asserting whether p, one must *in fact* believe whether p. A competent assertor can offer an assertion only if the assertion does in fact represent the beliefs of the assertor.

To clarify the differences between the strong requirement that one must in fact believe whether p, which is defended by myself, and the weaker requirement that it must be reasonable for one to believe whether p, which is defended by Lackey, consider the following modified version of the creationist teacher presented by Lackey:

Stella is a devoutly Christian fourth-grade teacher, and her religious beliefs are grounded in a deep faith that she has had since she was a very young child. Part of this faith includes a belief in the truth of creationism and, accordingly, a belief in the falsity of the evolutionary theory. Despite this, Stella fully recognizes that there is an overwhelming amount of scientific evidence against both of these beliefs. Indeed, she readily admits that she is not basing her own commitment to creationism on evidence at all but, rather, on the personal faith that she has in an all-powerful Creator. Because of this, Stella does not think that religion is something that she should impose on those around her, and this is especially true with respect to her fourth-grade students. Instead, she regards her duty as a teacher to include presenting material that is best supported by the available evidence, which clearly includes the truth of the evolutionary theory. As a result, while presenting her biology lesson today, Stella asserts to her students, “Modern-day Homo sapiens evolved from Homo erectus”, though she herself does not believe this proposition. [28, p. 111]

Stella⁸ has strong evidence that Homo sapiens evolved from Homo erectus, and she asserts this proposition to her students despite not actually believing it herself. In this case, Stella does not possess a belief in the proposition; nevertheless, she has substantial evidence indicating that such a proposition should be believed, making it reasonable for her to believe the proposition. However, Stella must have believed that p to genuinely assert that p, because competence in the realm of assertion, I argue, requires that one offer an assertion in the presence of the corresponding belief. In this regard, a strong requirement for being competent for an assertion is required. Fulfilling the competence norms requires a stronger epistemic condition than it being reasonable for a person to believe a proposition; one must actually believe a given proposition, and that belief must be justified. To qualify as competent in asserting a proposition, one must have a doxastic rather than a propositional justification for the given proposition. Respectively, in my view, Stella does not qualify as a person who is competent to assert whether Homo sapiens evolved from Homo erectus because she does not in fact believe the proposition. Hence, she violates a competence norm. Even if she had intended to speak justifiably, I think she would not have been competent to assert whether Homo sapiens evolved from Homo erectus.

The norm related to promise-making has been discussed: a competence norm. If an assertion is a matter of promising to speak justifiably regarding whether p, and simultaneously keeping or breaking that promise, we should expect it to be governed by the norm relevant to promise-making and by the norm relevant to promise-keeping. The norm related to promise-keeping is as follows:

- (b) asserting regarding whether p involves speaking justifiably or unjustifiably regarding whether p (i.e., keeping or breaking the promise).
- One must assert regarding whether p only if one does in fact speak justifiably regarding whether p.

Trusting other people’s words involves relying upon them to fulfil a commitment – to satisfy both promise-making and promise-keeping norms. A trustworthy assertor must conform to both promise-making and promise-keeping norms. A trustworthy assertor must:

- be competent to speak justifiably regarding whether p;
- only if one is ex post justified in believing whether p (constitutional competence);

Footnote 7 (continued)

tion), nor am I discussing under what conditions hearers are justified in believing what a speaker testifies. I simply clarify the condition required to be met for an assertor to be competent in asserting that p.

⁸ Those who argue for propositional justification need to demonstrate that Stella treats consideration, which plays the role of evidence, as evidence. If she did not see considerations as evidence, she would not have evidence. Accordingly, she would not be justified in the sense of proposition.

- only if one is awake and alert (conditional competence);
- only if one can ensure what audiences expect to hear from them (situational competence).
- in fact speak justifiably regarding whether p .

3 Artificial intelligence and trust

I have addressed the question, ‘How do A and B cultivate or maintain the relationship of trust?’, and discussed the norms of being trustworthy regarding a general task of ϕ and a specific task of assertion, emphasising the role of B as a trusted person in maintaining or cultivating a trust relationship with A. I now take the final step towards answering the main question: ‘How does an AI system affect privacy?’ This step requires exploring how an AI system impacts trust relationships to answer the following question: ‘How does an AI system affect trust relationships between A and B?’ Answering this question is essential to accomplishing the main goal of this research: understanding how an AI system impacts privacy, which depends upon trust.

How does an AI system affect trust relationships between A and B? To answer this, I examine how AI impacts B’s competence. The main case study of this paper is the one in which the assessor (B) employs an AI system to decide whether A has breast cancer (p). One norm an assessor must fulfil to be trustworthy is being competent to speak justifiably regarding whether p . In doing so, the assessor must be *ex post* justified in believing whether p . Therefore, the question that may arise is whether and how the assessor is justified, in a doxastic sense, in declaring whether p in cases in which p is a proposition delivered by an AI system. Part of the answer to this question emphasises the role of AI in justifying B’s belief that p , and thus its contribution to B’s competence.

When an AI system, as the diagnostic instrument, informs B that the scan or biopsy of the patient (A) indicates the presence of cancerous cells, B uses the instrument as an ‘epistemic instrument’ [29, p. 118] and forms beliefs based on what the instrument delivers, and then acts accordingly. Grindrod [30] refers to beliefs formed based on deliverance from an AI system in general, or an ML model in particular, as computational beliefs. ‘Whether’ and ‘how’ B are justified through believing the proposition delivered by the instrument. In other words, *how* is B’s computational belief justified?⁹

The question of how to justify B’s computational belief hinges on whether such a belief can be regarded as a distinctive form of belief or as an epistemic source that can be reduced to other epistemic sources. According to Goldman [26], a distinctive source provides justification on its own, without depending on other sources for its justificatory power, whereas reductionism-based justification is derived from other, more basic sources. In addition to memory and perception, I consider testimony as a distinct epistemic source. In line with Grindrod [30], I endorse the reductionism approach to computational belief, even though these types of belief cannot be reduced to memory, perception, and testimony. Rather, computational belief can be viewed as a form of inferential belief that acquires justificatory power from reliable inductive inference.

Computational beliefs cannot be reduced to memory, perception, or testimony. Memory can be dismissed because the process of obtaining a computational belief is not equivalent to remembering a certain proposition. Computational beliefs do not resemble perceptual beliefs either; perceptual experiences with an instrument justify B in believing merely that there is an instrument, rather than believing in that deliverance. As a result, computational beliefs are not completely captured as a form of perceptual belief. Computational beliefs cannot be described as the result of a testimonial exchange. An AI system is not an epistemic agent; it does not possess beliefs in the common sense. Therefore, we cannot rely upon an AI system via testimony [30].

I agree with Grindrod [30] that beliefs formed based on the deliverance of an AI system can be reduced to a form of inferential beliefs. B might infer computational belief that p from premises that take the form of inductive generalisation reasoning or, alternatively, premises that describe what other people testify to [29, 30]. Accordingly, B might apply at least two distinct arguments to explain how he reaches the conclusion that p . However, B is not obliged to offer A a justification for what is said, nor does B need to undertake justificatory responsibility for what he says (see Sect. 2.2.2). Since B’s doxastic attitude towards the proposition that p is justified only if arriving at the belief that p is the output of a reliable process (see Sect. 2.2.1), the justification of B’s belief in p that can be offered for each distinct argument is presented as follows.

First, B might reach his computational belief that p by appealing to premises that describe a merely observed correlation, which offers him inductive support for the target proposition p :

- P1: The deliverance of the instrument is proposition p .
- P2: B learns from experience and test data samples that the given instrument in this specific field usually delivers the correct proposition.

⁹ In what follows, I do not consider the ‘whether’ question. Readers wanting to know whether beliefs formed via the results of ML models can be justified at all should see Grindrod’s [30] paper, which provides reasons to think that computational beliefs are justified.

P3: The deliverance of propositions p by the given instrument in this specific field is correct.

Therefore,

C: p .

Suppose B uses the system with no particular view regarding its reliability. He uses the personal data of those whose diseases have not been diagnosed by the system as test data to assess the accuracy and model performance. He finds the system produces correct answers for the test data (P2) and eventually infers the deliverance of the system in this context (or specific field) is epistemically reliable (P3). Recall principle (R). This inductive inferential cognitive step involved a conditional reliable process. That is, the step's later output is usually true if the earlier input to it is true. Given that B's experiences with tested data are the input of the process, the output that is P3 is reliable. His belief in p was then formed using another inferential step, which is a conditionally reliable belief-forming process. B's belief in p is the output of the inferential process with the input of inductive generalisation. Since both reasoning processes are reliable, then, according to principle (R), B's belief in p is justified.

Second, B might infer some computational belief that p is based on premises upon which he relies regarding the testimony of another person:

P1: The deliverance of the instrument is proposition p .

P2: Other person said that the given instrument in this specific field usually delivers the correct proposition.

P3: The deliverance of proposition p by the given instrument in this specific field is correct.

Therefore,

C: p .

Again, according to principle (R), B is justified in believing p because p is the output of the inferential reasoning process, which is a conditionally reliable process. The input of this process is a testimonial belief (P2), which can itself be considered a conditionally reliable process or unconditionally reliable process. In the debate about testimonial knowledge, there has been a great deal of discussion about whether testimony as an epistemic source can be reduced to basic epistemic sources [11], or whether it constitutes a separate and distinct epistemic source [31]. Regarding the former, testimonial belief can be formed based on the process that is conditionally reliable with the input of memory, perceptual, or other inferential belief. Regarding the latter, testimonial belief is formed based on the process that is unconditionally reliable. Either way, the input of the process is reliable. Thus, B's belief in p is justified.¹⁰

¹⁰ In both arguments, B's belief in p is justified not by evidence—*beliefs* from which p can be inferred, or *perceptual* and *memory* expe-

I have explained that B is justified in believing that p because of the existence of a valid inferential process that forms this belief. In the first case, B relies on the inductive generalisation that proceeds from the limited sample of B's case to infer his belief in p . In the second case, B relies on another's testimony to infer his belief that p . Therefore, justification of the computational belief involves, *first*, B's or, *second*, the other's cognitive accomplishments. Furthermore, either B himself tests and gains inductive support for the accuracy of the instrument, or the developer of the ML model testifies to some level of accuracy for the model; therefore, *third*, B's computational beliefs partly rely on the accuracy and the operation of the instrument. Hence, in addition to B's or the other's cognitive accomplishments, this feature of the ML model's performance contributes to the justification of the computational belief.

First, concerning B's cognitive accomplishment, does it require that B be aware of how the instrument operates to be justified in believing that p ? Does the accomplishment require that B understands how the instrument he relies on performs to form a justified belief based on what the instrument delivers? The answer, in my view, is negative. According to the above discussion, being justified in believing that p is independent of being aware of how the instrument operates; rather, it requires that the belief is the output of a valid reasoning process in which the input beliefs are reliable. Although it is often not possible to understand properly how the algorithm processes the data and reaches the outcome it does, such an opacity does not impact the reasoning process that justifies a computational belief.

However, such an opacity leads to a significant issue, which is 'epistemic responsibility gaps' [30]. According to Grindrod [30], there is an important sense in which B relies on his epistemic community while employing instruments he does not understand. The epistemic community consists of individuals who comprehend how the instrument performs, and B can appeal to that community if they find that the instrumental inferences are incorrect. However, computational beliefs depend upon autonomous learning algorithms, which are opaque in nature, making it challenging for any member or group of members to understand the exact workings of these algorithms. Therefore, B cannot properly rely on his epistemic community to compensate for his not understanding how the instrument performs when he forms his computational belief [30].

Second, concerning the role of the other's cognitive accomplishment in justifying B's computational belief, does it require that the epistemic community be aware of how the

Footnote 10 (continued)

riences—but by non-evidentiary reasons concerning the reliability of the processes involved in forming p .

instrument operates to testify to the accuracy of the instrument? Do the responsibility gaps impact B's justification for believing what the instrument delivers? Again, in my view, the answer is negative. It is not necessary for the person who developed an instrument or model to understand how it operates to testify to its accuracy. Without necessarily understanding how the instrument operates, the model developer can appropriately declare that the instrument performs accurately as they have credence in the instrument's performance, which is supported by testing sample datasets. Therefore, a lack of epistemic responsibility by the epistemic community has no effect on the justification of the computational belief. The lack of impact does not imply the discussion of epistemic responsibility does not merit investigation. On the contrary, computational belief leads to a distinct structure of epistemic responsibility, which deserves detailed research, but not in the realm of appropriate assertion and trust.

Third, it is argued that B's computational beliefs partly rely on the accuracy and the operation of the instrument. Although a lack of understanding of how an AI system (or an ML model) performs does not affect the justification of the computational belief, its accuracy does. Since being justified in believing what the instrument delivers is required for B to be competent in what he asserts, the accuracy of an AI model affects B's competence. Given that trustworthiness requires competence, an AI system impacts trust relationships between A and B since B's competence requires the AI system to perform accurately.

4 Conclusion: trust, privacy, and artificial intelligence

How does an AI system affect privacy? This section summarises the previous discussions and answers this question. A person (B) who employs an AI system to respond to another person's (A) question (p) relies epistemically upon the system and asserts p based on what the system delivers. One norm that B must fulfil to be trustworthy is the competence to speak justifiably regarding whether p. Justification of B's belief that p partly relies on the accuracy of the AI system. Thus, accuracy is a feature of an AI system's performance that contributes to the justification of B's belief in p. Accordingly, B's competence relies on the accuracy of the operation of the system. Since trustworthiness requires B's competence while asserting p, the AI system affects trustworthiness and, consequently, the trust relationship between A and B.

Privacy is a social value constituted by trust-based relationships. Privacy, in a disclosure context, is constituted by interactions between different individuals based on trust. Since AI affects trust, AI impacts privacy. To achieve privacy as a social value, an AI system must perform accurately. Hence, the main RQ, concerning how an AI system

affects privacy, is explained by how an accurate AI system contributes to building trust relationships between A and B, which constitutes privacy. As a result, both B, as the trustee, and the AI system that makes B competent in his assertion contribute to the constituting of privacy.

To conclude, I believe, in contexts in which the relationships among individuals engaged in the practice of information-sharing are grounded in trust, that sharing information, analysing, and inferring from the shared information, as well as preserving privacy, are not mutually exclusive.

4.1 Implication: extending the scope of privacy

Does taking trust as an ex post approach impact the scope of privacy? To answer this question, it is crucial to study the type of information within the scope of privacy. According to Inness [32], privacy might not protect *all* information about a person, but might involve only *intimate* information. The intimacy of information stems from the act of sharing information that is itself intimate. An act or activity is intimate iff its meaning and values draw from the person's intimate motivations, such as love, liking, or care. The act of sharing information is intimate iff it is understood to take its meaning and value from our love, liking, or care, not merely if it conveys a desire on our part to inform another person. For example, we value showing our love letters to others as an intimate act iff it conveys the meaning that we care for them, not to extort money from them. Protecting privacy entails protecting actions (such as the dissemination of information about oneself) that are understood as expressions of love, liking, or care; privacy claims are claims to exercise control over intimate decisions and actions.

Inness's idea has two interrelated parts: the realm of privacy and privacy claims. Therefore, it is important to discuss how taking trust as an ex post approach to privacy affects these parts. I begin with the privacy realm part. By taking trust as an ex post approach to privacy, the scope of privacy is expanded to include information exchanged in a trust-based context. Unlike intimate relationships formed between friends, partners, and lovers, trust relationships are not always confined to those people who know them and are close to them. Although trust does not require a person to be in a close relationship, it subsumes cases in which the person is in an intimate relationship. In this regard, the scope of privacy is expanded to include information shared or revealed in a trust-based context.

Determining the scope of privacy does not require a perspectival assessment, because assessing trust does not demand a perspectival assessment. Unlike intimacy, which requires a personal viewpoint to characterise underlying motivations—a person can confirm whether their own actions embody love, liking, or care—interpersonal trust is independent of one's motivation. A person motivated

to act is not trustworthy; rather, trustworthiness requires avoiding unfulfilled commitments or broken promises (see Sect. 2.2.1).

Regarding privacy claims, unlike Inness's taking control account of privacy, which merely emphasises a person who shares data with others, the trust-based approach emphasises the role of others and relations between them in constituting privacy as well. Privacy claims are claims that the information exchanged in the trust-based context is to be cared for. Such a claim can take the form of cultivating trust between those involved in a disclosure context by conforming to trust norms. Accordingly, protecting privacy entails promoting or maintaining trust. Therefore, regulations need to be established that focus on building, maintaining, and fostering trust in a disclosure context.

Acknowledgements I would like to extend my special thanks to Dr. Adam Henschke for his advice and valuable comments on each version of the manuscript, which significantly contributed to enhancing its quality. Additionally, I am grateful to Dr. Björn Lundgren for his constructive feedback on the initial draft of the manuscript, which played a crucial role in improving its overall quality.

Author contributions HA is the single contributor of this paper.

Funding Haleh Asgarinia is supported by the Ph.D. fellowship from the 'PROTECT- Protecting Personal Data Amidst Big Data Innovation' project, funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813497.

Availability of data and material Not applicable.

Code availability Not applicable.

Declarations

Conflict of interest The author has no competing interests to declare. The paper in part or in full has not been submitted or published anywhere. The paper will not be submitted elsewhere until the editorial process is completed.

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Steeves, V.: Reclaiming the social value of privacy. In: Kerr, I., Steeves, V., Lucock, C. (eds.) *Lessons from the Identity Trail: Anonymity, Privacy, and Identity in a Networked Society*, pp. 109–208. Oxford University Press, New York (2009)
2. Altman, I.: A conceptual analysis. *Environ. Behav.* **8**(1), 1 (1976). <https://doi.org/10.1177/001391657600800102>
3. Altman, I.: *The Environment and Social Behavior: Privacy, Personal Space, Territory*. Brooks/Cole Publishing Company, Crowding (1975)
4. Nissenbaum, H.F.: *Privacy in context: technology, policy, and the integrity of social life*. Stanford Law Books, Stanford, Calif (2010)
5. Waldman, A.E.: Privacy as Trust: Sharing Personal Information in a Networked World. *Univ. Miami Law Rev.* **69**, 559 (2015)
6. Waldman, A.E.: *Privacy as Trust*. Cambridge University Press (2018)
7. Richards, N.M., Hartzog, W.: A relational turn for data protection? *Eur. Data Prot. Law Rev.* **6**(4), 492–497 (2020). <https://doi.org/10.21552/edpl/2020/4/5>
8. Akselrod-Ballin, A., et al.: Predicting breast cancer by applying deep learning to linked health records and mammograms. *Radiology* **292**(2), 331–342 (2019). <https://doi.org/10.1148/radiol.2019182622>
9. Hawley, K.: *How To Be Trustworthy* | Oxford Academic. Oxford University Press (2019)
10. Baier, A.: Trust and antitrust. *Ethics* **96**(2), 231–260 (1986). <https://doi.org/10.1086/292745>
11. Hardin, R.: Trust and trustworthiness. In: *Trust and trustworthiness*. Russell Sage Foundation, New York, NY, US (2002)
12. Hawley, K.: Trust, distrust and commitment. *Noûs* **48**(1), 1–20 (2014). <https://doi.org/10.1111/nous.12000>
13. Hieronymi, P.: The reasons of trust. *Australas. J. Philos.* **86**(2), 213–236 (2008). <https://doi.org/10.1080/00048400801886496>
14. Holton, R.: Deciding to trust, coming to believe. *Australas. J. Philos.* **72**(1), 63–76 (1994). <https://doi.org/10.1080/00048409412345881>
15. Fricker, E.: Trust and Testimonial Justification, presented at the Midwest Epistemology Conference, Notre Dame, 1–19 (2018).
16. Carter, J. A., and Simion, M.: The ethics and epistemology of trust. *Internet Encyclopedia of Philosophy* (2021).
17. Faulkner, P.: The attitude of trust is basic. *Analysis* **75**(3), 424–429 (2015). <https://doi.org/10.1093/analys/anv037>
18. Jones, K.: Trust as an affective attitude. *Ethics* **107**(1), 4–25 (1996)
19. Faulkner, P.: The problem of trust. In: Faulkner, P., Simpson, T. (eds.) *The Philosophy of Trust*. Oxford University Press (2017)
20. Hinchman, E.S.: Telling as inviting to trust. *Philos. Phenomenol. Res.* **70**(3), 562–587 (2005). <https://doi.org/10.1111/j.1933-1592.2005.tb00415.x>
21. McLeod, C.: Trust. In: Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab Stanford University (2021)
22. Strawson, P. F.: Freedom and resentment, in *Freedom and resentment, and other essays*, First Edition., London: Methuen [distributed in the USA by Harper & Row, Barnes & Noble Import Division, 1–25 (1974).
23. Sosa, E.: How competence matters in epistemology. *Philos. Perspect.* **24**(1), 465–475 (2010). <https://doi.org/10.1111/j.1520-8583.2010.00200.x>
24. Goldman, A.I.: What is justified belief? In: Pappas, G.S. (ed.) *Justification and Knowledge: New Studies in Epistemology*, pp. 1–23. Springer Netherlands, Dordrecht (1979)

25. Goldman, A.I.: The structure of justification. In: Goldman, A.I., McGrath, M. (eds.) *Epistemology: A Contemporary Introduction*, pp. 3–24. Oxford University Press (2015)
26. Goldman, A.I.: Two debates about justification: evidentialism vs. reliabilism and internalism vs. externalism. In: Goldman, A.I., McGrath, M. (eds.) *Epistemology: A Contemporary Introduction*, pp. 25–50. Oxford University Press (2015)
27. Brandom, R.: Asserting. *Noûs* **17**(4), 637–650 (1983). <https://doi.org/10.2307/2215086>
28. Lackey, J.: *Learning From Words: Testimony as a Source of Knowledge*. Oxford University Press, Oxford (2008)
29. Sosa, E.: Knowledge: instrumental and testimonial. In: Lackey, J., Sosa, E. (eds.) *The Epistemology of Testimony*, pp. 116–124. Oxford University Press (2006). <https://doi.org/10.1093/acprof:oso/9780199276011.003.0006>
30. Grindrod, J.: Computational beliefs. *Inquiry* (2019). <https://doi.org/10.1080/0020174X.2019.1688178>
31. Coady, C.A.J.: Testimony and observation. *Am. Philos. Q.* **10**(2), 149–155 (1973)
32. Inness, J.C.: *Privacy, Intimacy, and Isolation*. Oxford University Press (1992)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.