

Human-centric Literature on Trust for SfTI Veracity Spearhead

Kelly Blincoe¹, Markus Luczak-Roesch², Tim Miller², and Matthias Galster³

¹The University of Auckland, New Zealand, k.blincoe@auckland.ac.nz

²Victoria University of Wellington, New Zealand, markus.luczak-roesch@vuw.ac.nz,
tim.miller@vuw.ac.nz

³University of Canterbury, New Zealand, matthias.galster@canterbury.ac.nz

Abstract

This article summarizes the literature on trust of digital technologies from a human-centric perspective. We summarize literature on trust in face-to-face interactions from other fields, followed by a discussion of organizational trust, technology-mediated trust, trust of software products, trust of AI, and blockchain. This report was created for the Science for Technological Innovation Veracity Spearhead supported by New Zealand's National Science Challenges.

1 Trust in General

Disclaimer: This report does not include a Māori perspective on trust. Work is in progress to combine this Western view with a Māori perspective.

Trust has been described as facilitating cooperative behaviour [1]. Trust has been examined extensively in the fields of experimental psychology, philosophy, sociology, and political science [2]. In sociology, trust is considered to be multi-faceted with distinct cognitive, emotional, and behavioral dimensions [2]. The cognitive dimension says that trust is based on rational decisions, while the emotional dimension, which is also referred to as affect-based trust, says that emotional relationships between people form a basis for trust [3, 4]. The behavioral component of trust is the action of doing something with uncertain outcomes while assuming that all people involved in the action will act with integrity [5]. Trust is only needed when there is some level of risk or uncertainty involved, so trust also involves vulnerability [6]. Research in psychology and philosophy also describe trust as having both rational and emotional aspects [7].

Much research has considered how initial trust is formed. McKnight and Chervany identified a set of characteristics based on a review of literature across multiple disciplines (see Figure 1) [8]. First, a person's own disposition to trust, or their willingness or tendency to depend on others, impacts how trust is formed. Second, there must be conditions in place that could lead to success,

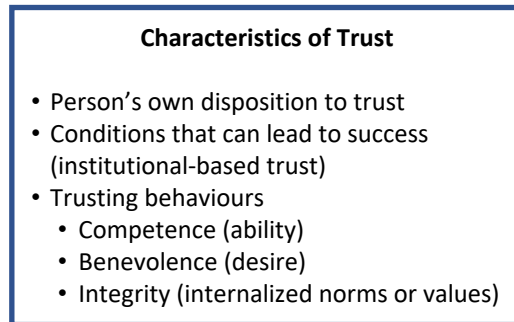


Figure 1: Characteristics of Trust [9]

this is called institutional-based trust. Finally, there are three main trusting behaviours: competence, benevolence, and integrity. Competence is defined as the belief that the other party has the required skills, benevolence is the belief that the other party wants to do good, and integrity relates to the belief that the other party has good values or character.

Studies have shown that people are influenced by a “truth bias”, meaning that they are more likely to assess things as truth than lies, even when being deceived [10, 11, 12]. People are not very good at identifying deception, which is defined as someone purposely misleading someone else [13]. Research in psychology has found that people assume all information as truth initially and only later change their assessment if they find the information is false [14]. This is in line with the Truth-Default theory, which states that people tend to believe each other by default [15].

When trust is damaged, there are negative consequences [16, 17]. Researchers have studied how trust is repaired after it has been damaged [18, 19]. Unlike initial trust, significant effort is often required to rebuild trust after a trust violation. Various factors impact trust repair, including the strength of the initial trust [20]. A theory which is important for trust repair is attribution theory [21]. Attribution theory considers the cause of the trust violation. It considers three main dimensions: locus of control (internal or external), controllability, and stability [19]. Attribution theory states that the outcomes and reactions of trust violations will vary based on these dimensions. The theory also suggests that the outcomes and reactions are not permanent and that trust can be repaired following violations [22].

2 Organizational Trust

While most research in the field of psychology has focused on interpersonal trust, organizational trust has been studied extensively in other domains like management and marketing [6]. Organizational trust has been defined as “the belief that the decision makers will produce outcomes favorable to the person’s interests without any influence by the person” [23]. Management researchers argue that trust can improve business performance [6, 24]. Trust within an organization can result in improved productivity and satisfaction of employees [25]. In the field of marketing, researchers have found that consumers’ trust of a business is impacted by both the people within that business that they interact with and the business’s management practices and policies [26].

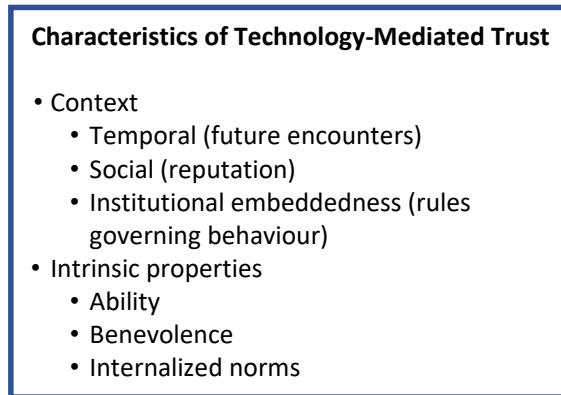


Figure 2: Characteristics of trust from Riegelsberger et al. framework for technology-mediated interactions [9]

3 Technology-Mediated Trust

Trust has primarily been studied from a perspective of human, face-to-face interactions [27]. When interactions occur through technology, signals of trust are different [1]. Riegelsberger et al. proposed a framework of trust in technology-mediated interactions (see Figure 2), which included both contextual and intrinsic properties of trust [9]. The contextual properties included in the framework are temporal, social, and institutional embeddedness. Where temporal embeddedness considers likely future encounters since repeated interactions can both encourage trustworthy behaviour and provide signals to make decisions around trust. Social embeddedness considers the reputation of the person or organization who is being trusted, which can be discussed and shared across the trustors (those doing the trusting). Institutional embeddedness refers to the institutions that govern behaviour, such as judicial systems or organizations, since the rules imposed by these institutions will influence trust. Yet, there is an acknowledgement that new technology can disrupt trust formation, since new technology has the potential to transform the way in which people interact, which can lead to uncertainty and vulnerability until new norms are established [28].

The intrinsic properties included in Riegelsberger et al.’s framework are ability, internalized norms, and benevolence, which are in line with the trusting behaviours of competence, integrity, and benevolence of McKnight and Chervany described above [8]. Here, ability refers to the capabilities and characteristics of the person or organization who is being trusted that will enable them to fulfill the promised outcomes. Internalized norms includes attributes such as honesty, credibility, reliability, dependability, openness, and good will. Benevolence represents the enjoyment obtained by person or organization who is being trusted when good outcomes are experienced by the person doing the trusting.

While this framework was created to conceptualise trust in technology-mediated interactions, the elements of trust are still focused on the people or organizations involved in the trust relationship. Shneiderman also recognized this in his definition of trust: “If users rely on a computer and it fails, they may get frustrated or vent their anger by smashing a keyboard, but there is no relationship of trust with a computer. If users depend on a network and it breaks, they cannot get compensation

from the network. However, they can seek compensation from people or organizations they trusted to supply a correctly functioning computer or communication service.” [1] Based on this definition, Shneiderman developed a set of guidelines for developers of online services, such as e-commerce or e-services, which are underlined by two key principles. First, the organization providing the service should ensure trust both by providing evidence of past trustworthy performance and providing strong assurances of trust. Second, the organization should clarify responsibilities and obligations by providing full disclosure of terms, guarantees, and mechanisms for disputes.

Of course, when considering technology-mediated trust, it is also important to know that people have different perceptions of and attitudes towards technology, and so technology-mediated trust will be subjective [29]. This is in line with the characteristics of trust in general by McKnight and Chervany which state that a person’s own disposition to trust impacts trust formation.

4 Trust of Software

In line with this, studies have shown that for software products, trust is based on both a trust of the creators of the software product and a trust of the software itself [30, 31]. Similarly, Siau and Wang argue that trust in technology is determined by three main factors: human characteristics, environment characteristics, and technology characteristics [32]. Users of software products assess the trustworthiness of software in different ways [31, 33, 34]. Yang et al. propose a software trust framework which considers software correctness, security, and reliability as measures of trustworthiness [33]. Jackson equates trust to dependability of the software product to perform a particular task [31]. These definitions all relate to the intrinsic ability and internalized norms of the software. On the other hand, Wang et al. show that user feedback of software products is useful to determining levels of trust, which relates to the reputation of the software [35]. Mercuri considered the view of transparency as it relates to trust of software, defining different ways that transparency can be achieved [36]. For example, through open sourced code, certifications, and assurances. Provenance, defined as “metadata about the origin, context or history of data”, can also promote transparency [37].

Another perspective comes from the field of human computer interaction where the relationship between trust and user interface design has been studied. Interface designers argue that the visual design of the interface forms the first impressions of trust [38]. Rendell et al. found that inclusion of nature imagery on websites positively influenced users’ perceptions of trust [39]. Xiling found that simple and well laid out interfaces promoted trust [40]. They also found that familiarity, being able to clearly relate the “offline” brand and experience to the online interface, was important for trust. Xiling found that usability was important for building trust [40]. Systems that were easy to use, consistent, and logically structured were more trusted. While researchers have investigated the use of particular colors in an interface and their relationship with user trust, no relationship was identified [41].

5 Trust of AI

With the rise of Artificial Intelligence (AI) to perform decision-making, it is also important to consider trust as it relates to these systems in particular. Glikson and Woolley find that the

representation of an AI system (e.g., robot, humanoid, embedded) and the system’s capabilities are important factors in developing trust [42]. Siau and Wang present a list of factors that are used for both building initial trust in AI systems and developing continuous trust in those systems [32]. They find that understanding how AI works (its transparency and explainability) and being able to trial the AI system before adopting it (trialability) are important for initial trust formation. In addition, the visual appearance of the AI (its representation), reviews of the AI system written by other users, and the users’ perceptions of AI in general based on exposure to things like media coverage or Sci-fi books will also impact initial trust.

While many AI systems operate as black boxes (there is no way to understand why decisions are being made), transparency, explainability, and interpretability are still seen as important for trust of AI systems [43]. Some research defines interpretable as understandable or transparent [44]. Others define interpretable as providing explanations for decisions. In these cases, the model may not be transparent, but some understandable reasons for decisions are provided by the black box AI model [45]. Thus, interpretability may be defined as both explainable or transparent. Explanations are often proposed to improve trust in AI systems [46, 47] and recent research shows that software users do want explanations when complex decisions are being made [48].

For developing continuous trust in AI systems, Siau and Wang [32] find that usability and reliability, collaboration and communication, sociability and bonding, security and privacy protection, interpretability, concerns about job replacement, and goal congruence are important factors. Of course, accuracy is also important. Yin et al. found that people considered both a model’s stated accuracy and its observed accuracy in determining their trust of the model [49]. Siau and Wang say “trust in AI takes time to build, seconds to break, and forever to repair once it is broken!” [32]

We have seen many examples where AI has gone wrong, and Winfield and Jirotko argue that ethical governance is critical to building trust in AI [50]. Through a literature review of trust and AI, Lockey et al. identified five main challenges: 1) transparency and explainability, 2) accuracy and reliability, 3) automation resulting in job loss, 4) anthropomorphism (or including human-like characteristics) leading to over-estimation of the AI system, and 5) privacy concerns related to mass data extraction [51].

Another important factor related to trust in AI is fairness [52]. While one might assume machines can make more fair decisions that are free from human bias, it is well known that AI systems actually amplify existing bias [53, 54]. Historical bias in training data can cause AI systems to learn this bias and make biased decisions.

Accountability is also important [52] for trust in AI. This factor considers who will be held responsible for the decisions made by AI systems. There is currently not a clear answer to who should be held accountable. The Law Commission of England and Wales and the Scottish Law Commission recently proposed that self driving car users should not be held responsible for crashes and other driving offenses.¹ However, in the US, self driving car users are considered responsible. Research suggests more auditing of AI is needed to reduce corporate reputation damage and assure AI is legal, ethical, and safe [55].

Trust of AI also comes down to perceptions of how decisions are made. Machines make decisions which are rule-based and algorithmic [56]. Machines do not consider emotions in their decision-making nor can they learn in the same way as humans [57]. These differences can lead to algorithmic aversion, where people prefer human made decisions even if the decisions are inferior to those made

¹<https://www.forbes.com/sites/zacharysmith/2022/01/25/self-driving-car-users-shouldnt-be-held-responsible-for-crashes-uk-report-says>

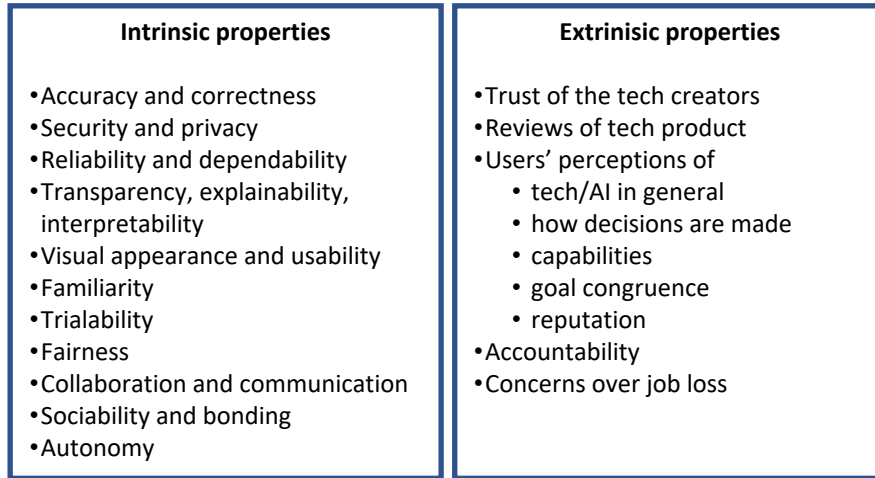


Figure 3: Factors that affect trust of digital technology

by a machine [58].

Jussupow et al. defined four characteristics of algorithms that influence aversion: 1) algorithm agency which describes the level at which the algorithm behaves autonomously; 2) algorithm performance which considers the accuracy and failures of the algorithm; 3) perceived algorithm capabilities which describes the algorithm’s perceived ability to perform the task; and 4) human involvement which relates to how much humans (but not the end user) are involved in training and using the algorithm [58].

Recent research has found that people do not want AI to make moral decisions [59]. Through a series of studies, Bigman and Gray found that people distrust AI to make moral decisions even when the outcome is favorable since “machines can neither fully think nor feel” [59]. They suggest that limiting AI to providing only advice and increasing the AI’s perceived experience and expertise are ways to improve trust in AI for moral decisions [59].

Figure 3 summarizes the factors that influence trust for digital technologies, including software products and AI.

6 Blockchain

A discussion on trust of software is not complete without a mention of Blockchain. Blockchain has been nicknamed a “trustless” technology [60, 61]. It emerged due to a growing lack of trust in centralized systems which relied on trust of institutions (e.g. banks) [62]. The idea of a trustless technology goes back to the theory of Wang and Emurian that claims that trust is only needed when there is some level of risk or uncertainty involved [6]. Blockchain is a distributed, immutable ledger. This means transactions cannot be modified once they are written to the ledger and all participants have access to the shared ledger. Thus, “users subject themselves to the authority of a technological system that they are confident is immutable, rather than to the authority of centralized institutions which are deemed untrustworthy.” [62]

De Filippi et al. prefer to label Blockchains as “confidence machines”, since their underlying technology creates shared expectations and confidence in the correctness of its transactions. However, while many Blockchains remove the need to trust a single organization, they still require “distributed trust” since there are often a large number of actors who require a low-level of trust [62]. There still needs to be trust that the data going into the Blockchain can be trusted since compromised data cannot be corrected. These actors must be trusted not to collude and cause collective harm. It should also be noted that not all Blockchains follow this same model and some (e.g. The Linux Foundation’s Hyperledger and Amazon’s QLDB) are maintained by organizations, which means these organizations will also still require trust. Using blockchain may lead to trade-offs between trust and other concerns like energy consumption and sustainability [63]. These tradeoffs could in turn compromise trust because benevolence and integrity can be adversely affected if a Blockchain is perceived to be non-sustainable.

There does not appear to be literature on trust violations and repair in relation to blockchain technology.

References

- [1] B. Shneiderman, “Designing trust into online experiences,” *Communications of the ACM*, vol. 43, no. 12, pp. 57–59, 2000.
- [2] J. D. Lewis and A. Weigert, “Trust as a social reality,” *Social forces*, vol. 63, no. 4, pp. 967–985, 1985.
- [3] D. J. McAllister, “Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations,” *Academy of management journal*, vol. 38, no. 1, pp. 24–59, 1995.
- [4] S. Chowdhury, “The role of affect-and cognition-based trust in complex knowledge sharing,” *Journal of Managerial issues*, pp. 310–326, 2005.
- [5] B. Barber, *The logic and limits of trust*. Rutgers University Press, 1983.
- [6] Y. D. Wang and H. H. Emurian, “An overview of online trust: Concepts, elements, and implications,” *Computers in human behavior*, vol. 21, no. 1, pp. 105–125, 2005.
- [7] D. Trček, “A brief overview of trust and reputation over various domains,” *Trust and Reputation Management Systems*, pp. 5–19, 2018.
- [8] D. H. McKnight and N. L. Chervany, “What trust means in e-commerce customer relationships: An interdisciplinary conceptual typology,” *International journal of electronic commerce*, vol. 6, no. 2, pp. 35–59, 2001.
- [9] J. Riegelsberger, M. A. Sasse, and J. D. McCarthy, “The mechanics of trust: A framework for research and design,” *International Journal of Human-Computer Studies*, vol. 62, no. 3, pp. 381–422, 2005.
- [10] S. A. McCornack and M. R. Parks, “Deception detection and relationship development: The other side of trust,” *Annals of the International Communication Association*, vol. 9, no. 1, pp. 377–389, 1986.

- [11] M. Zuckerman, B. M. DePaulo, and R. Rosenthal, "Verbal and nonverbal communication of deception," in *Advances in experimental social psychology*, vol. 14, pp. 1–59, Elsevier, 1981.
- [12] M. Zuckerman, R. Koestner, M. J. Colella, and A. O. Alton, "Anchoring in the detection of deception and leakage.," *Journal of Personality and Social Psychology*, vol. 47, no. 2, p. 301, 1984.
- [13] T. R. Levine, H. S. Park, and S. A. McCornack, "Accuracy in detecting truths and lies: Documenting the "veracity effect",," *Communications Monographs*, vol. 66, no. 2, pp. 125–144, 1999.
- [14] D. T. Gilbert, D. S. Krull, and P. S. Malone, "Unbelieving the unbelievable: Some problems in the rejection of false information.,," *Journal of personality and social psychology*, vol. 59, no. 4, p. 601, 1990.
- [15] T. R. Levine, "Truth-default theory (tdt) a theory of human deception and deception detection," *Journal of Language and Social Psychology*, vol. 33, no. 4, pp. 378–392, 2014.
- [16] R. J. Lewicki, B. B. Bunker, *et al.*, "Developing and maintaining trust in work relationships," *Trust in organizations: Frontiers of theory and research*, vol. 114, p. 139, 1996.
- [17] S. L. Robinson, "Trust and breach of the psychological contract," *Administrative science quarterly*, pp. 574–599, 1996.
- [18] P. H. Kim, K. T. Dirks, and C. D. Cooper, "The repair of trust: A dynamic bilateral perspective and multilevel conceptualization," *Academy of Management Review*, vol. 34, no. 3, pp. 401–422, 2009.
- [19] E. C. Tomlinson and R. C. Mryer, "The role of causal attribution dimensions in trust repair," *Academy of management review*, vol. 34, no. 1, pp. 85–104, 2009.
- [20] R. J. Lewicki and C. Wiethoff, "Trust, trust development, and trust repair," *The handbook of conflict resolution: Theory and practice*, vol. 1, no. 1, pp. 86–107, 2000.
- [21] B. Weiner, "An attributional theory of achievement motivation and emotion.,," *Psychological review*, vol. 92, no. 4, p. 548, 1985.
- [22] G. Bansal and F. M. Zahedi, "Trust violation and repair: The information privacy perspective," *Decision Support Systems*, vol. 71, pp. 62–77, 2015.
- [23] J. W. Driscoll, "Trust and participation in organizational decision making as predictors of satisfaction," *Academy of management journal*, vol. 21, no. 1, pp. 44–56, 1978.
- [24] M. Sako *et al.*, "Does trust improve business performance," *Organisafional trust: A reader*, pp. 267–294, 2006.
- [25] R. C. Mayer, J. H. Davis, and F. D. Schoorman, "An integrative model of organizational trust," *Academy of management review*, vol. 20, no. 3, pp. 709–734, 1995.
- [26] D. Sirdeshmukh, J. Singh, and B. Sabol, "Consumer trust, value, and loyalty in relational exchanges," *Journal of marketing*, vol. 66, no. 1, pp. 15–37, 2002.

- [27] D. H. McKnight and N. L. Chervany, “The meanings of trust,” tech. rep., Carlson School of Management, University of Minnesota, 1996.
- [28] D. H. McKnight and N. L. Chervany, “What is trust? a conceptual analysis and an interdisciplinary model,” in *Americas Conference on Information Systems*, 2000.
- [29] S. Grabner-Kraeuter, “The role of consumers’ trust in online-shopping,” *Journal of Business Ethics*, vol. 39, no. 1, pp. 43–50, 2002.
- [30] M. Söllner, A. Hoffmann, and J. M. Leimeister, “Why different trust relationships matter for information systems users,” *European Journal of Information Systems*, vol. 25, no. 3, pp. 274–287, 2016.
- [31] D. Jackson, “A direct path to dependable software,” *Communications of the ACM*, vol. 52, no. 4, pp. 78–88, 2009.
- [32] K. Siau and W. Wang, “Building trust in artificial intelligence, machine learning, and robotics,” *Cutter business technology journal*, vol. 31, no. 2, pp. 47–53, 2018.
- [33] X. Yang, G. Jabeen, P. Luo, X.-L. Zhu, and M.-H. Liu, “A unified measurement solution of software trustworthiness based on social-to-software framework,” *Journal of Computer Science and Technology*, vol. 33, no. 3, pp. 603–620, 2018.
- [34] F. S. Grodzinsky, K. W. Miller, and M. J. Wolf, “Developing artificial agents worthy of trust: “would you buy a used car from this artificial agent?”,” *Ethics and information technology*, vol. 13, no. 1, pp. 17–27, 2011.
- [35] B. Wang, Y. Chen, S. Zhang, and H. Wu, “Updating model of software component trustworthiness based on users feedback,” *IEEE Access*, vol. 7, pp. 60199–60205, 2019.
- [36] R. T. Mercuri, “Trusting in transparency,” *Communications of the ACM*, vol. 48, no. 5, pp. 15–19, 2005.
- [37] J. Cheney, S. Chong, N. Foster, M. Seltzer, and S. Vansummeren, “Provenance: a future history,” in *Proceedings of the 24th ACM SIGPLAN conference companion on Object oriented programming systems languages and applications*, pp. 957–964, 2009.
- [38] S. Weinschenk, *100 things every designer needs to know about people*. Pearson Education, 2011.
- [39] A. Rendell, M. T. Adam, A. Eidels, and T. Teubner, “Nature imagery in user interface design: the influence on user perceptions of trust and aesthetics,” *Behaviour & Information Technology*, pp. 1–17, 2021.
- [40] Z. Xiling and L. Xiangchun, “Effective user interface design for consumer trust: Two case studies,” 2005.
- [41] F. Hawlitschek, L.-E. Jansen, E. Lux, T. Teubner, and C. Weinhardt, “Colors and trust: The influence of user interface design on trust and reciprocity,” in *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pp. 590–599, IEEE, 2016.
- [42] E. Glikson and A. W. Woolley, “Human trust in artificial intelligence: Review of empirical research,” *Academy of Management Annals*, vol. 14, no. 2, pp. 627–660, 2020.

- [43] Z. C. Lipton, “The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery.,” *Queue*, vol. 16, no. 3, pp. 31–57, 2018.
- [44] Y. Lou, R. Caruana, J. Gehrke, and G. Hooker, “Accurate intelligible models with pairwise interactions,” in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 623–631, 2013.
- [45] Y. Lou, R. Caruana, and J. Gehrke, “Intelligible models for classification and regression,” in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 150–158, 2012.
- [46] B. Y. Lim, Q. Yang, A. M. Abdul, and D. Wang, “Why these explanations? selecting intelligibility types for explanation goals.,” in *IUI Workshops*, 2019.
- [47] M. Sadeghi, V. Klös, and A. Vogelsang, “Cases for explainable software systems: Characteristics and examples,” in *2021 IEEE 29th International Requirements Engineering Conference Workshops (REW)*, pp. 181–187, IEEE, 2021.
- [48] L. Chazette and K. Schneider, “Explainability as a non-functional requirement: challenges and recommendations,” *Requirements Engineering*, vol. 25, no. 4, pp. 493–514, 2020.
- [49] M. Yin, J. Wortman Vaughan, and H. Wallach, “Understanding the effect of accuracy on trust in machine learning models,” in *Proceedings of the 2019 chi conference on human factors in computing systems*, pp. 1–12, 2019.
- [50] A. F. Winfield and M. Jirotko, “Ethical governance is essential to building trust in robotics and artificial intelligence systems,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 376, no. 2133, p. 20180085, 2018.
- [51] S. Lockey, N. Gillespie, D. Holm, and I. A. Someh, “A review of trust in artificial intelligence: challenges, vulnerabilities and future directions,” in *Proceedings of the Annual Hawaii International Conference on System Sciences*, vol. 2020, pp. 5463–5472, Hawaii International Conference on System Sciences, 2021.
- [52] V. Vakkuri, K.-K. Kemell, J. Kultanen, and P. Abrahamsson, “The current state of industrial practice in artificial intelligence ethics,” *IEEE Software*, vol. 37, no. 4, pp. 50–57, 2020.
- [53] P. S. Chauhan and N. Kshetri, “The role of data and artificial intelligence in driving diversity, equity, and inclusion,” *Computer*, vol. 55, no. 4, pp. 88–93, 2022.
- [54] E. Ntoutsis, P. Fafalios, U. Gadiraju, V. Iosifidis, W. Nejdl, M.-E. Vidal, S. Ruggieri, F. Turini, S. Papadopoulos, E. Krasanakis, *et al.*, “Bias in data-driven artificial intelligence systems—an introductory survey,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 10, no. 3, p. e1356, 2020.
- [55] A. Koshiyama, E. Kazim, and P. Treleaven, “Algorithm auditing: Managing the legal, ethical, and technological risks of artificial intelligence, machine learning, and associated algorithms,” *Computer*, vol. 55, no. 4, pp. 40–50, 2022.
- [56] B. J. Dietvorst, J. P. Simmons, and C. Massey, “Algorithm aversion: people erroneously avoid algorithms after seeing them err.,” *Journal of Experimental Psychology: General*, vol. 144, no. 1, p. 114, 2015.

- [57] F. Cuzzolin, A. Morelli, B. Cirstea, and B. J. Sahakian, “Knowing me, knowing you: theory of mind in ai,” *Psychological medicine*, vol. 50, no. 7, pp. 1057–1061, 2020.
- [58] E. Jussupow, I. Benbasat, and A. Heinzl, “Why are we averse towards algorithms? a comprehensive literature review on algorithm aversion,” in *European Conference on Information Systems*, 2020.
- [59] Y. E. Bigman and K. Gray, “People are averse to machines making moral decisions,” *Cognition*, vol. 181, pp. 21–34, 2018.
- [60] G. Vidan and V. Lehdonvirta, “Mine the gap: Bitcoin and the maintenance of trustlessness,” *New Media & Society*, vol. 21, no. 1, pp. 42–59, 2019.
- [61] F. Hawlitschek, B. Notheisen, and T. Teubner, “The limits of trust-free systems: A literature review on blockchain technology and trust in the sharing economy,” *Electronic commerce research and applications*, vol. 29, pp. 50–63, 2018.
- [62] P. De Filippi, M. Mannan, and W. Reijers, “Blockchain as a confidence machine: The problem of trust & challenges of governance,” *Technology in Society*, vol. 62, p. 101284, 2020.
- [63] J. Sedlmeir, H. U. Buhl, G. Fridgen, and R. Keller, “The energy consumption of blockchain technology: beyond myth,” *Business & Information Systems Engineering*, vol. 62, no. 6, pp. 599–608, 2020.