Robots in Community

Willow Wong and Mark Findlay

***Abstract*—Building trust between humans and robots involves more than improving the technical reliability or safety of the machine. Without a proper way to address and resolve the deep-seated human fear of being replaced by robots in situations of essential human engagements, the machine will continue to take the blame. This paper proposes an alternative approach to trust-building, where practitioners, policymakers, and regulators work together to engage members of the community in deciding how increasingly smart robots should be designed and used to improve their communal living.**

***Keywords*—social robots; acceptability and trust; design and human factors; emotional robotics; ethics and philosophy.**

I. Introduction

Human-Robot Interactions (HRI) is an emerging field of study that integrates interdisciplinary perspectives into the engineering project of creating robots with desired attributes (e.g., safety, reliability, transparency, friendliness and trustworthiness). More specifically, trusted robotics [1] aims at harmonizing the interactions between the human user and the robots to shape the overall trajectory of human-robot relations. However, if treated only at the superficial level of apparent interaction, the likelihood of fostering sustainable trusted human-robot relationships, wherein robots assume a more socially embedded role in communities, becomes unworkable. This brief paper explores the limitations of the current approach and offers an alternative position of trust-building as a community-based practice that stimulates and sustains harmonious bonding between humans and robots.

II. Trust As A Technical Challenge?

"Do you trust politicians? Do you trust teachers? If somebody said to you, [...] do you trust elementary school teachers? You would probably begin by saying, 'To do what?' [...] And you might say, when you understood the answer to that, "Well, I trust some of them, but not others." That's a perfectly rational thing. […] I would aim to have more trust in the trustworthy but not in the untrustworthy. In fact, I aim positively to try not to trust the untrustworthy." [2].

Although there is no single definition of trust, O'Neil's view above differentiating trust and trustworthiness is a good starting point: 'trustworthiness' as a judgement drawn from other technical qualities about the robot's competency at performing its intended tasks; whereas 'trust' is a human response – emotive *and* rational in nature – to the presence of these technical qualities.

This project is supported by the National Research Foundation, Singapore under its Emerging Areas Research Projects (EARP) Funding Initiative. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore.

The authors are with the Centre for AI and Data Governance, Yong Pung How School of Law, Singapore Management University (e-mail: willowwong@smu.edu.sg; markfindlay@smu.edu.sg).

In ideal scenarios, the people and tools we choose to trust will be trustworthy. But assertions of the robot's competency and efficacy does not automatically translate trustworthiness to user trust. This gap may be due to the *static* technical aspects of robot systems, which do not capture the dynamic and *shifting* concerns of human trust. That is, while machine trustworthiness is measured within laboratory settings, human trust exists in complex social environments, wherein changes in people, places and external circumstances may trigger corresponding shifts in how trust is manifested within social relations [3].

Robot competencies constitute an *essential* foundation to trust-building between humans and robots. After all, it is difficult for the user to have confidence in the robot's embedded role in the community if its basic performance is not reliable or safe. However, these are not *only* requirement for trust to occur. Human trust has always been connected with a deep-rooted sense of vulnerability [4-5]. When deploying robots into the community, practitioners, policymakers and regulators will need to be mindful of potential triggers of social insecurities and distrust that could work against the early formations of trusted human-robot relations (as will be discussed further in Section III). In the context of robotics, the aim to build trusted human-robot relations will likely involve more than the governance of the robot's technical system.

Drawing from [6], the limitations of approaching human trust as a technical challenge can be summarized as such:

A. Limits of Blind Trust

The mainstream governance approach in Singapore aims to create trust between humans and machines by drawing heavily on abstract ethical principles and safety certificates to co-deliver the quality of 'trustworthiness' within technological systems [7]. But the average user cannot recognize or have meaningful engagements with the merits of these frameworks without the relevant expertise or knowledge. In this scenario, knowledge gap engenders trust gap. Users who feel discouraged may then 'switch off' and default to exercising blind trust in the decision-making actors and processes [8]. The manifestation of blind trust, where user do not know why they have chosen to trust, may not provide momentum necessary to maintain human-robot trust relations that are sustained from one encounter to the next.

B. Limits of Rationality

Users' initial perceptions and subsequent judgements of a robot's trustworthiness are not strictly rational actions. The human choice between trust and mistrust are often instinctive, habitual, and motivated by emotions [5-6]. Even when trust has been extended, it can be revoked or suspended because of the users' pre-existing suspicions or reservations separate to the encounter in question.

For example, the Covid-19 vaccination has been increasingly enforced as a necessity – without which one would face barriers in accessing dining venues, international travel and job opportunities [9-11]. Some have extended conditional trust to the vaccination for the ease of continuing with their everyday living. But for others, this incursion to their bodily autonomy [12] has triggered trust erosion towards scientists, governments and other institutions enforcing vaccination uptake to 'solve' the global pandemic.

Whether robots and other technologies (e.g. artificial intelligence) are seen as a threat depends on the human user's wider life experiences of trust relations, which inform their evolving personal understanding of what *feels* dangerous, foreign, or intrusive. Thus, rational consideration of the robot's technical safety is not the only factor influencing the users' thought processes.

C. *Limits of Objectivity*

Robots, as technological artifacts, inherit the existing statuses of trust *and* mistrust that human users extend towards the brand or maker of the robots. When deployed by a prominent source of authority (e.g. a governmental agency), user receptivity towards the robots and their intended purposes reveals the common attitude towards the powers using these technologies. In this sense, user perceptions of the robot's trustworthiness are grounded by communal contexts; what trusted humans-robots relations means in practical terms will also be shaped by social and cultural norms. As such, 'trustworthiness' as a quality cannot be a completely value-neutral or objective attribute.

These limitations clarify that human trust is heavily influenced by *perceptions*, rather than actual technical knowledge of the robot's safety and capabilities. But this dimension is not well-addressed by existing efforts from key decision-makers to defend the robot's trustworthiness. While developers and practitioners typically focus on improving the robot's functionality, regulators and policy makers implement frameworks that attribute the integrity of technical systems to the structured of governance [13]. But community voices have minimal presence in this decision-making frame.

Implicit in this technical approach is the assumption that human trust will arise as the natural outcome of achieving sufficient levels of technical robustness. If there are external factors obstructing trust formation between the human and the robot, concerns of individuals and the wider community are treated as secondary and separate to the scope of the technical design (if not dismissed completely). But human trust does not behave as a passive target that marks the end of the practitioners or regulators' efforts in 'fixing' the robot.

Even though human drivers are not perfect drivers, many see driverless cars as riskier (regardless of the actual technical specifications). Negative perceptions are intensified by news reports of fatal accidents caused by cases of Tesla drivers overestimating the capabilities of the driver-assistance system to navigate roads and avoid potential collisions [14]. The prevention of *over-trust* will depend on safer design and active community education to minimize misunderstanding of what driverless cars can and cannot do [15].

Assurances that a machine can work in a predictable and safe manner only hold weight when the human recipient recognizes the legitimacy and authority of those extending assurances. The social dimensions of trust formation requires a more holistic understanding of what 'trustworthy' means to the average human being (outside of expert and regulatory frames of reference). If trust manifests as social relationships, ongoing maintenance and negotiation of these trusted relationships will be necessary to build healthy connections among key decision parties in the robotics ecosystem.

We argue that trust is simply a *signal* – which can shift over time – of the relative strength of participatory technology development processes. Using trust and distrust as a signal, decision-makers can undertake community engagements to map out what tends to trigger frightful and insecure (or otherwise, open and accepting) responses towards the design and roll-out of robotics applications in the community. Trust becomes the emergent property of trusted relationships flourishing within community spaces.

If the scope of work carved out above seem radically different and distant from the developers' principal task of ensuring the robot's performance is safe and reliable (and therefore, trustworthy), it is important remember that our goal to improve user acceptability is built on the foundations of trusted *human-robot* relations. The 'human' dimension of trust-building must be treated with the same, if not more, priority as the 'robot' aspect of the same equation. Indeed, modelling trusting or distrusting human interactions outside of controlled laboratories settings can also reveal conditions that work against or in favour of human-robot trust relations.

As such, the core argument of this paper is that approaching trust-building as a community-based practice can strengthen the receptivity of the best engineering projects. Laying down strong foundations for community engagements (as complementary to technical considerations of robot safety) will activate meaningful and sustainable human-robot interactions and bring to fruition the engineer's aspirations to apply robotics for social good. If community perceptions of 'trustworthiness' receives the same research energy as creating 'trustworthy' robot design, the functionality of robots will likely be integrated into shared community spaces.

III. Alternative Approaches

Drawing inspiration from [15-16], this paper proposes a community-focused approach to build trust between users and key decision-makers in the wider robotics ecosystems; this groundwork is *prior* and foundational to human-robot trust. The ethics of trusted robotics is a decision-making framework geared towards trust outcomes by aligning the emerging capabilities of robotics with the evolving needs and priorities of human recipients across diverse communities.

Since trust is shaped by contextual elements, such as social and cultural norms, we argue there is no one-size-fits-all approach to design and deploy robots operating with the optimal or desired levels of human compatibility. We should exercise great caution in adopting a top-down approach of design, creation and production that may transplant ethical standards and values originating from the Global North into localized communities of developing countries from the Global South [17]. This reservation may seem less significant in the globalized contexts of Singapore, but if there is a desire in government thinking to position the nation as an international hub for cutting-edge technology [18], regional thinking and local sentiments should not be overlooked.

Small details such as the visual appearances of the robots, or the accents of their speech which do not resonate with certain cultures, religions, traditions and communities, risk alienation rather than integration. It is also vital to recognize that ethical principles have specific cultural locations and do not exist in a vacuum. In Asian societies, the value of the 'communal' self [19] may translate well extensive communal dialogues about where or how robots ought to belong in the community. This topic can still gain traction in Western societies due to privacy implications of data collected by robotics technology [20]. The comon goal of human-robot trusted relations may signal different meaning and aspirations across diverse communities.

For this reason, we posit the ideal pathway of achieving 'trusted robotics' is whichever systems, or modes, of ethical decision-making that can most effectively stimulate positive trust outcomes and human receptivity towards robots. This will likely involve participatory pathways to leverage local voices and prioritize local knowledge in decision-making. Arnstein's Ladder of Citizen Participation reveal the necessity within community engagements to redistribute powers via negotiation between citizens (i.e. human recipients of robotics technology) and powerholders (i.e. those responsible for robot design, manufacturing and deployment) [21]. This will ensure ethical decision-making of trusted robotics move beyond the tokenistic modes of participation. Similarly, the Democracy Cube by Archon Fung illustrates diverse modes of public participation within democratic decision-making frames [22].

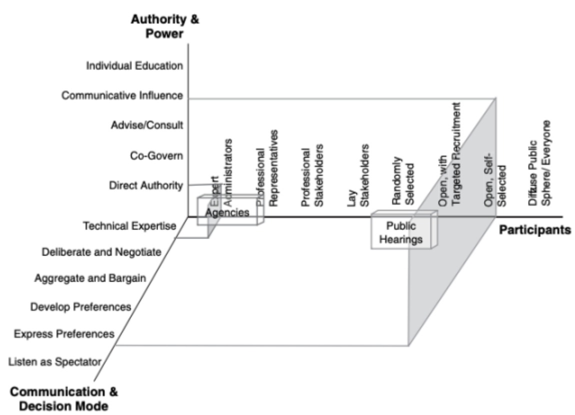

Figure 1. Fung's Democracy Cube integrates three spectrums across three axes: communication & decision modes, authority & power and participant types.

A good starting point to implement participatory pathways is a small-scale focus group consultation. It has been shown to address and placate the concerns of therapists' on the appropriateness and safety to use social robots in therapy settings [23]. Referrals from the qualified clinicians and medical professionals who trust the competencies of the robot can also influence later adoptions by other colleagues in the medical field, alongside better receptivity from the patients [24] – thereby revealing robot-human trust as built upon the conditions of patient-clinician trust.

One may still question whether this mode of small-scale engagement can be scaled to integrate the voices of the wider community members. After all, it may be too time-consuming and impractical to seek direct engagements with every member of the community for every potential robotics roll-out project. In this case, leveraging online modes of participation will enable the public to vocalize their concerns for the adoption of technological systems.

For example, an online nationwide survey done in Germany to understand public acceptance of digital farming technologies (DFT) has engaged a representative sample of the German population based on gender, age, education and size of place of residence [25]. When treated as a starting point – rather than a tokenistic consultation exercise – to activate trust in the community, representative surveys can yield insights relevant for practitioners, policymakers, and regulators to gauge current trust landscapes. If increasing use of DFT is inevitable, methodological reflections on whose voices were not captured in the survey will guide future efforts in engaging the public with the deliberative process of co-creating the intended design, purposes, and usages of robots in community spaces.

In the case of Tai O, a fishing town in Hong Kong, the data research team held meetings with community leaders to identify the potential for ethical surveillance to address the 'pain point' of the community (i.e. heavy pedestrian traffic from tourists) [26]. The team installed low-impact mobility monitors under the active guidance of local contacts, whose local knowledge enabled relevant data to be captured. The data insights were used to further the community's ongoing policy advocacy efforts. The positive community reception of ethical surveillance roll-out at Tai O was due to the collaborative nature of the decision-making process. This is contrasted against the installation of 'smart lampposts' by the government, which were torn down by residents due to privacy concerns [27]. Even with the safest lamppost design and guarantees that the data collected will not be misused by the government, the community still rejected the smart lampposts' purposes and roll-out locations. In this case, the lack of prior community input led to negative public perception, as uninvited and intrusive machine presence in community spaces were met with distrust and hostility.

In the contexts of robotics roll-out, community engagement is crucial to identify key social issues that needs addressing – if the purpose receives widespread endorsement by the public, the applications of robots to resolve community pain-points will also likely receive positive reception. This is especially true when community leaders, officials and engineers work together to co-create the design, intended utility and roll-out method of robots. In this picture, accessible language plays a key role. Effective modes of communication can help connect the public, who may be unfamiliar with technical jargons, with key parties in the robotics ecosystem. It is by establishing a common understanding that the public can then embrace *why* and *how* robotics can be used to achieve social good in their communities.

Contrary to the trending understanding of 'trusted robotics' as branding technological artifacts or their creators with a fundamentally *static* 'trustworthy' certification [29], we have proposed trust as a positive signal of active engagements in our communal duty to ensure emerging technologies are used to achieve social good. If practitioners, policymakers and regulators can dispel negative perceptions of the robots' threat, community participation into decision-making processes will likely arise from a place of openness, as opposed to defensive and hostile fear. Trust building and maintenance, in this sense, is a dynamic communal practice, akin to a lived and *evolving* phenomenon manifesting within social life-spaces.

IV. On Emotionality

A field that runs adjacent to trusted robotics is affective robotics, aiming to develop the robot' ability to detect, analyze, and *mirror* human-like emotions or social behaviours [30-32]. This method of instilling or validating user trust towards robots is different from trusted robotics' focus to strive at trust by aligning the safe developments of robot technologies with the needs of human recipients across diverse communities.

Drawing from section II (A-C), the causal influences of trust gaps reveal the *double-edged reality* of emotion-driven perceptions as the trigger of trust and distrust. On the one hand, affective robotics may be crucial in 'increasing a robot's ability to increase its cooperative movements and display mutual supportiveness, (which) will enhance the user's sense of commitment and therefore, trust, in the robot' [3]. On the other hand, using emotions as the *primary* device to actualize HRI may increase the risks of humanoid robot exacerbating existing tensions between humans and machines.

Although robots with human-like characteristics (e.g. the ability to use empathetic speech patterns to match the emotional state of the user) can elicit higher levels of trust in a social setting [33-34], they can provoke higher levels of user disappointment and distrust if their behaviours do not conform to user expectations [35]. For example, sentiments of distrust may become intensified in rehabilitation settings as vulnerable users may perceive a 'friendly' socially assistive robots as manipulative and disingenuous [36]. Although the claim that the user's mental and emotional states can be 'detected' has been disputed [37], user may have inflated expectations of human-like robots' ability to 'mindread' their needs. This may damage to trust relations in the long run, as errors arising from unsafe human-robot interactions risks encouraging a 'blame the robot' attitude (as an extension of existing 'blame the algorithm' sentiments) [38-39].

Our interest in the question of 'to anthropomorphize or not' is not to dispute the value of robots that are more capable of performing key duties because they have, as an example, similar range of motion as the human hand. Rather, our concern is the alleged compatibility between human emotions and the long-term sustainability of trust-building between humans and machines. Without consistent effort to maintain trust relations, emotive trust can deteriorate and easily shift into distrust. It is risky to assume robots mirroring human expressions and emotive speech can be a safe shortcut to harmonious human-robot relations.

In rehabilitation settings, trust-building requires more than a smiley face on the robot, as actual trust-repair depends on the social robot's ability to acknowledge and apologize for their errors [4]. Thus, whether human trust evoked by affective robotics techniques is proportionate *and* constructive towards the long-term trajectory of harmonious human-robot relations still remains unanswered.

If the powerful influences of human emotion cannot be easily harnessed without consequences (e.g. unintended social harm or damaging in human-robot trust relations), affective robotics should not be used as the primary method to build human-robot trust. Indeed, the foundational belief that "the more human-like the robots, the better they can facilitate trust" will also need further exploration and justification.

V. Conclusion

Building trust between humans and robots involves more than the technical challenge of 'injecting' emotions or 'trustworthiness' inside the machine. We do not doubt that existing commitments to improve the technical robustness of robotic systems will contribute to the foundations of human-robot trust relations. However, approach trust-building as a community practice can enhance the outcome of the best engineering projects. Without a proper way to address and resolve the deep-seated human fear of being replaced by robots in situations of essential human engagements, the machine will continue to take the blame.

Emphasis on fixing the technical system misses the key to building human-robot trust: genuine efforts in raising public awareness of the various compliance metrics of 'trustworthy' or 'ethical' robots cannot be a truly meaningful exercise, if no prior communal dialogue has occurred for people – with diverse sets of needs, priorities and aspirations – to input into early design formulations or subsequent methods of deployments into community spaces. To this end, certification that a robot can be trusted effectively outsources the personal and relationship-oriented dimensions of trust, and thus, may have little impact in retaining trust after the piece of technology has been rolled out into the community.

As trust does not reside as a static quality within the machine, rather it manifests as a lived and communal phenomenon, the exchange of ideas on what trust cannot be done by elites or experts alone. By initiating a public conversation about the intended purposes and benefits of robotics, especially when linked with the use of AI and big data, practitioners, policymakers, and regulators can better decide how increasingly smart robots should be designed and used to improve their communal living.

Ultimately, we believe that the ingredients that engender trust may be found partially in laboratories, but even more so in community-centric engagements that can activate trust to flow through communal relations of people constituting the wider decision-making ecosystem.

Acknowledgment

We thank Assistant Professor Harold Soh, Department of Computer Science at the National University of Singapore (NUS) for his generous contributions to intellectually stimulating conversations on trusted robotics.

REFERENCES

[1] B. C. Kok and H. Soh, “Trust in Robots: Challenges and Opportunities,” *Current Robotics Reports*, vol. 1, no. 4, pp. 297–309, Dec. 2020, doi: 10.1007/s43154-020-00029-y.

[2] O. O’Neill, *What we don’t understand about trust*, (Jun. 2013). Accessed: Sep. 12, 2021. [Online Video]. Available: https://www.ted.com/talks/onora_o_neill_what_we_don_t_understand_about_trust

[3] R. Cotterrell, “Law, Emotion and Affective Community,” *Social Science Research Network*, Rochester, NY, SSRN Scholarly Paper ID 3212860, Jul. 2018. Accessed: Dec. 03, 2020. [Online]. Available: https://papers.ssrn.com/abstract=3212860

[4] A. Langer, R. Feingold-Polak, O. Mueller, P. Kellmeyer, and S. Levy-Tzedek, “Trust in socially assistive robots: Considerations for use in rehabilitation,” *Neuroscience & Biobehavioral Reviews*, vol. 104, pp. 231–239, Sep. 2019, doi: 10.1016/j.neubiorev.2019.07.014.

[5] A. Fabris, “Can We Trust Machines? The Role of Trust in Technological Environments,” in Trust, Cham, 2020, pp. 123–135. doi: 10.1007/978-3-030-44018-3_9.

[6] M. Findlay and W. Wong, “Trust and Regulation: An Analysis of Emotion,” *Social Science Research Network*, Rochester, NY, SSRN Scholarly Paper ID 3857447, Jun. 2021. doi: 10.2139/ssrn.3857447.

[7] R. Kurohi, “S’pore to invest $50m over 5 years to build digital trust,” *The Straits Times*, Singapore, Jul. 15, 2021. Accessed: Aug. 13, 2021. [Online]. Available: https://www.straitstimes.com/tech/spore-to-invest-50m-over-5-years-to-build-digital-trust

[8] M. Taddeo, “On The Risks of Trusting Artificial Intelligence: The Case of Cybersecurity,” *Social Science Research Network*, Rochester, NY, SSRN Scholarly Paper ID 3730651, Nov. 2020. doi: 10.2139/ssrn.3730651.

[9] “Los Angeles makes Covid-19 vaccines compulsory for schoolkids,” *The Straits Times*, Singapore, Sep. 10, 2021. Accessed: Sep. 13, 2021. [Online]. Available: https://www.straitstimes.com/world/united-states/los-angeles-makes-covid-19-vaccines-compulsory-for-schoolkids

[10] M. Shields and B. H. Neghaiwi, “Swiss will need COVID certificates to go to bars, restaurants,” *Reuters*, Sep. 08, 2021. Accessed: Sep. 13, 2021. [Online]. Available: https://www.reuters.com/world/europe/swiss-cabinet-tightens-coronavirus-curbs-protect-hospitals-2021-09-08/

[11] J. Karl, “How Far Are the Unvaccinated Willing to Go?,” *Bloomberg.com*, Sep. 12, 2021. Accessed: Sep. 13, 2021. [Online]. Available: https://www.bloomberg.com/opinion/articles/2021-09-12/covid-19-vaccine-s-merits-are-simple-and-effective-kth60xpt

[12] T. Hals, “Factbox: COVID-19 and the U.S. courts: challenges to vaccine requirements,” Reuters, Aug. 06, 2021. Accessed: Sep. 13, 2021. [Online]. Available: https://www.reuters.com/world/us/covid-19-us-courts-challenges-vaccine-requirements-2021-08-06/

[13] B. W. Israelsen and N. R. Ahmed, “‘Dave...I can assure you ...that it’s going to be all right ...’ A Definition, Case for, and Survey of Algorithmic Assurances in Human-Autonomy Trust Relationships,” *ACM Comput. Surv.*, vol. 51, no. 6, pp. 1–37, Feb. 2019, doi: 10.1145/3267338.

[14] N. E. Boudette, “‘It Happened So Fast’: Inside a Fatal Tesla Autopilot Accident,” The New York Times, Aug. 17, 2021. Accessed: Sep. 13, 2021. [Online]. Available: https://www.nytimes.com/2021/08/17/business/tesla-autopilot-accident.html

[15] K. Holländer, P. Wintersberger, and A. Butz, “Overtrust in External Cues of Automated Vehicles: An Experimental Investigation,” in *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, New York, NY, USA, Sep. 2019, pp. 211–221. doi: 10.1145/3342197.3344528.

[16] M. Findlay and J. Seah, “An Ecosystem Approach to Ethical AI and Data Use: Experimental Reflections,” *arXiv:2101.02008 [cs]*, Dec. 2020, Accessed: Sep. 13, 2021. [Online]. Available: http://arxiv.org/abs/2101.02008

[17] J. Seah and M. Findlay, “Communicating Ethics across the AI Ecosystem,” *Social Science Research Network*, Rochester, NY, SSRN Scholarly Paper ID 3895522, Jul. 2021. doi: 10.2139/ssrn.3895522.

[18] H. Widdows, “Is Global Ethics Moral Neo-Colonialism? An Investigation of the Issue in the Context of Bioethics,” *Bioethics*, vol. 21, no. 6, pp. 305–315, 2007, doi: https://doi.org/10.1111/j.1467-8519.2007.00558.x.

[19] K. W. Wong and T. Bunnell, “‘New Economy’ Discourse and Spaces in Singapore: A Case Study of One-North,” *Environment and Planning A: Economy and Space*, vol. 38, no. 1, pp. 69–83, Jan. 2006, doi: 10.1068/a37229.

[20] W. T. D. Bary, *Asian Values and Human Rights: A Confucian Communitarian Perspective*. Harvard University Press, 1998.

[21] C. Lutz, M. Schöttler, and C. P. Hoffmann, “The privacy implications of social robots: Scoping review and expert interviews,” *Mobile Media & Communication*, vol. 7, no. 3, pp. 412–434, Sep. 2019, doi: 10.1177/2050157919843961.

[22] S. R. Arnstein, “A Ladder of Citizen Participation,” *Journal of the American Planning Association*, vol. 85, no. 1, pp. 24–34, Jan. 2019, doi: 10.1080/01944363.2018.1559388.

[23] A. Fung, “Varieties of Participation in Complex Governance,” *Public Administration Review*, vol. 66, no. s1, pp. 66–75, 2006, doi: 10.1111/j.1540-6210.2006.00667.x.

[24] K. Winkle, P. Caleb-Solly, A. Turton, and P. Bremner, “Social Robots for Engagement in Rehabilitative Therapies: Design Implications from a Study with Therapists,” in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, New York, NY, USA, Feb. 2018, pp. 289–297. doi: 10.1145/3171221.3171273.

[25] A. Alaiad and L. Zhou, “The determinants of home healthcare robots adoption: An empirical investigation,” *International Journal of Medical Informatics*, vol. 83, no. 11, pp. 825–840, Nov. 2014, doi: 10.1016/j.ijmedinf.2014.07.003.

[26] J. Pfeiffer, A. Gabriel, and M. Gandorfer, “Understanding the public attitudinal acceptance of digital farming technologies: a nationwide survey in Germany,” *Agriculture and Human Values*, vol. 38, no. 1, pp. 107–128, Feb. 2021, doi: 10.1007/s10460-020-10145-2.

[27] X. L. Wang, D. Elkin, and K. N. Chan, “Ethical Surveillance: Developing smart city surveillance through civic participation,” presented at the Ethics in AI Research Initiative for the Asia Pacific Roundtable, Sep. 10, 2021.

[28] E. Binder, “Hong Kong Protesters Spy a New Enemy: Lampposts,” *Wall Street Journal*, Aug. 30, 2019. Accessed: Sep. 13, 2021. [Online]. Available: https://www.wsj.com/articles/hong-kong-protesters-spy-a-new-enemy-lampposts-11567161002

[29] I. Tham, “Singapore’s private sector launches document for ethical use of AI; certification course to follow next year,” *The Straits Times*, Singapore, Oct. 16, 2020. Accessed: Aug. 13, 2021. [Online]. Available: https://www.straitstimes.com/tech/singapores-private-sector-launches-document-for-ethical-use-of-ai-certification-course-to

[30] K. Shaw, “Affectiva, SoftBank Robotics Team Up to Broaden Pepper’s Emotional Intelligence,” Robotics Business Review, Aug. 28, 2018. Accessed: Sep. 13, 2021. [Online]. Available: https://www.roboticsbusinessreview.com/news/emotional-intelligence-affectiva-softbank-robotics-team-up/

[31] R. Kirby, J. Forlizzi, and R. Simmons, “Affective social robots,” *Robotics and Autonomous Systems*, vol. 58, no. 3, pp. 322–332, Mar. 2010, doi: 10.1016/j.robot.2009.09.015.

[32] J. P. Sullins, “Friends by Design: A Design Philosophy for Personal Robotics Technology,” in *Philosophy and Design: From Engineering to Architecture*, P. Kroes, P. E. Vermaas, A. Light, and S. A. Moore, Eds. Dordrecht: Springer Netherlands, 2008, pp. 143–157. doi: 10.1007/978-1-4020-6591-0_11.

[33] A. Tapus, M. Maja, and B. Scassellatti, “The Grand Challenges in Socially Assistive Robotics,” *IEEE Robotics and Automation Magazine*, vol. 14, no. 1, p. N/A, 2007.

[34] M. Lewis, K. Sycara, and P. Walker, “The Role of Trust in Human-Robot Interaction,” in *Foundations of Trusted Autonomy*, H. A. Abbass, J. Scholz, and D. J. Reid, Eds. Cham: Springer International Publishing, 2018, pp. 135–159. doi: 10.1007/978-3-319-64816-3_8.

[35] J. Welge and M. Hassenzahl, "Better Than Human: About the Psychological Superpowers of Robots," in *Social Robotics*, Cham, 2016, pp. 993–1002. doi: 10.1007/978-3-319-47437-3_97.

[36] T. Vandemeulebroucke, B. D. de Casterlé, and C. Gastmans, "How do older adults experience and perceive socially assistive robots in aged care: a systematic review of qualitative evidence," *Aging & Mental Health*, vol. 22, no. 2, pp. 149–167, Feb. 2018, doi: 10.1080/13607863.2017.1286455.

[37] J. Wright, "Suspect AI: Vibraimage, Emotion Recognition Technology and Algorithmic Opacity," *Science, Technology and Society*, p. 1–20, Apr. 2021, doi: 10.1177/09717218211003411.

[38] S. Olhede and P. J. Wolfe, "Blame the algorithm?," *Significance*, vol. 17, no. 5, pp. 12–12, 2020, doi: 10.1111/1740-9713.01441.

[39] J. Möller, D. Trilling, N. Helberger, and B. van Es, "Do not blame it on the algorithm: an empirical assessment of multiple recommender systems and their impact on content diversity," Information, Communication & Society, vol. 21, no. 7, pp. 959–977, Jul. 2018, doi: 10.1080/1369118X.2018.1444076.