Human-Centered Design of Social Robots: Addressing Stakeholder Involvement to Enhance Long-Term Interaction Success

Andrew Blair, Mary Ellen Foster and Peggy Gregory

School of Computing Science, University of Glasgow, United Kingdom

Abstract

Real-world deployments of social robots are increasing, but the interactions often are gratuitous and do not last past an initial novelty effect. We propose that a major part of this is due to oversights within the design process, rather than necessarily a result of technical failures. We frame our argument within the domain of public spaces, one of the most complex environments for any robot. We highlight and discuss three important parts of any robotics system development; who we are considering and consulting to design the system, what we are asking them to contribute and be responsible for, and why this is necessary within human-robot interaction design.

Keywords

Social robotics, Stakeholder theory, Human-centered design, Co-design

1. Introduction

Real-world deployments of social robots have been increasing in recent years, with robots being found in shopping malls [1], exhibition spaces [2], train stations [3] and retail stores [4]. However, despite these appearances of robots in the wild, it is often found that a robot does not solve its intended use case. For example, in a care home in Japan, a survey revealed *"staff stopped using Hug after only a few days, saying it was cumbersome and time consuming to wheel from room to room—cutting into the time they had to interact with the residents"* [5]. This is often simply seen as an almost inevitable result due to "technical issues". However, we can see chatbots and even voice assistants deployed into live environments [6] successfully, taking advantage of the robustness and flexibility provided by modern Large Language Models. So, even though an embodied robot has exponentially more technical parts than a voice interface, is there still something else that the field is missing when deploying such robots in real-world contexts?

We argue that an important issue is poorly defined scope, whereby roboticists often deploy a robot to do a job that is too complex, inappropriate, or simply not useful—and when it fails, the robot lies unused. But why is the scope badly defined? As they are not generally domain experts of the deployment contexts, roboticists tend to make decisions based on technical convenience, or may misunderstand the domain and devise a solution that solves a problem that does not exist. Co-design is increasingly used within HRI to attempt to mitigate this recurrent issue [7, 8], but often this consists of a small group of stakeholders in one or two focus group sessions at the start of development. This is not sufficient; stakeholder engagement should be a continuous process with a wide range of stakeholders.

In this paper, we claim that, now that we have the technology to support resilient deployments, we need to expand *who* we are involving in robot system design, *what* these stakeholders are expected to contribute, and *why* this will allow long-term, successful robot deployments.

Socially Interactive Agent Applications, September 19, Glasgow, UK, 2024
Andrew.Blair@glasgow.ac.uk (A. Blair);
MaryEllen.Foster@glasgow.ac.uk (M. E. Foster);
Peggy.Gregory@glasgow.ac.uk (P. Gregory)
0000-0002-0453-1381 (A. Blair); 0000-0002-1228-7657 (M. E. Foster);
0000-0001-7891-6666 (P. Gregory)

2. Who?

In the context of project management, Freeman defined a stakeholder as *"any group or individual who can affect or is affected by the achievement of the organisation's objectives"* [9]. In the scope of an embodied robot system designed to operate in a public space, the size of the stakeholder list grows quickly. Stakeholders may be internal or external to the organisation, may be expert or novice users when it comes to social robotics, and may belong to groups such as external activists, trade unions, or even saboteurs. Correct stakeholder identification is critical to success; it makes sure the right people are engaged within a project, and therefore minimises the risk of failure. Techniques to identify stakeholders include brainstorming, snowball sampling [10] and using established baselines [11].

After defining the list of stakeholders, we can look to the salience model proposed by Mitchell et al. [12] for categorising and thus prioritising our stakeholders. In this model, stakeholders are rated based on three attributes; power, urgency, and legitimacy. Power is considered the influence a stakeholder has on the project; urgency is the degree to which immediate responses and/or action is required to a stakeholder's requirements; while legitimacy assesses a stakeholder's "right" to be involved in the process. This is an improvement over other models such as the power-interest matrix [13], where legitimacy and urgency are not considered, and as such many external stakeholders could be excluded from the outset. In the salience model (Figure 1), this results in eight distinct groupings, with one being non-stakeholders, where each should be engaged and managed in their own unique way.

If a stakeholder possesses all three attributes, then they are considered *Definitive* stakeholders; these are the stakeholders that have been primarily considered within the HRI design space. These stakeholders are often constitute three or four main groups depending on the deployment location. If it is within a lab space, it can be categorised as: the roboticists themselves, the subject matter experts, and the end users who will directly interact with the system itself. When it becomes a real-world deployment, often the management of a location will also become involved. These definitive stakeholders are, correctly, very important in any real-world deployment, as if any one of these stakeholders is not on board the deployment will likely stall before it even begins.

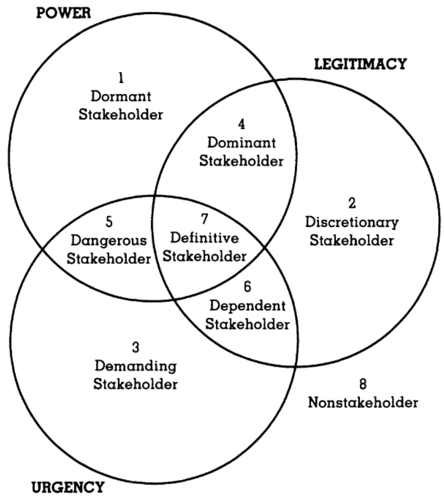

Figure 1: Stakeholder salience model [12]

However, the sphere of stakeholders, as we identified earlier, is much larger than simply definitive stakeholders. A small group of HRI researchers have begun recognising the importance of including these wider stakeholders, with a workshop on worker-robot relations [14] highlighting many of these, but often in isolation from each other; an overall approach is still missing from any robotic development project.

Dependent stakeholders, those with urgency and legitimacy, but no organisational power, are often key to understanding how a robot should actually operate on a day-to-day basis. In the case of public spaces, front-line workers are the largest group within this stakeholder definition. Whilst overall goals may be that the robot should help with a given domain, such as providing technical help, the reality can differ: for example, providing directions to the toilets ends up being more of an immediate customer need [15].

Dominant stakeholders, those with power and legitimacy, also need to be considered and engaged. This is where trade unions sit, who are definitely relevant especially in the case of larger-scale robotic deployments.

One group of stakeholders that need to be considered especially carefully is that of *Dangerous* stakeholders; the saboteurs are part of this group, holding power and urgency. Most public space robotic deployments consist of one, or at most a few, robots that cost a large amount of money to purchase and are hard, or even impossible to replace given funding constraints if anything happened to them. A dangerous stakeholder could be an activist, or someone who simply would attack the robot for amusement. Mitigation strategies include effectively employing researchers or front-line staff as not only technical support but "security guards" for the robot to ensure nothing untoward happens to it, or for the robot itself to be designed to move away from potentially dangerous users [16].

Dormant stakeholders are those considered to have high power but not actively involved in the project, such as director-level management of another part of the organisation. Dormant stakeholders have the power to cause the robotic deployment to be stopped entirely, but they do not have the time to attend every meeting, so they must be engaged at a level that allows them to be assured the deployment will cause no harm to organisations' reputation.

Discretionary stakeholders, those with high legitimacy, should be engaged where appropriate. This can often be charities or community groups, who may feel unsure about a robot being brought into a local space to them. Social responsibility is important, especially when bringing in often misunderstood technology to a space, and transparency about the goals of a deployment should be clearly communicated not only on request but also proactively.

Demanding stakeholders are those with high urgency. They require to be considered, for example a bystander asking questions about the robot but who is not a user of the robot, but do not generally pose a threat in the same way a Dangerous stakeholder does.

Generally, these three latter groups of stakeholders can be considered through scenario planning, and traditional risk management strategies. It may seem at first as if this is overplanning, but it is well established that the earlier challenges are identified the easier and cheaper they are to fix [17], and as such the benefits greatly outweigh any slight loss in efficiency.

3. What?

The **customer**, that is, who the robot is for, should always be the stakeholder responsible for the requirements, not the roboticists. However, robots are a new phenomenon for organisations, and as such understanding of them is likely to be limited. Roboticists need to be aware of this, or run the risk of the valuable stakeholder time running amok to Parkinson's Law of Triviality [18]. This satirical argument presents the idea that members of an organisation will discuss trivial matters at much greater lengths than the major decisions, as the trivial issues are much easier to grasp. This is something that can often be seen in software development projects, with the more informal term "bikeshedding" used to describe these discussions in open-source software. Robotics is no different; we can see exactly the same effect taking place, with participants talking about whether or not the robot should wear a hat, rather than how the robot should handle sensitive queries [19]. Proper meeting facilitation techniques, such as agendas and chairing, should be used, but roboticists should be prepared to answer questions on robot capabilities quickly and be able to distill them into a language that makes innate sense to meeting participants. If roboticists fail to do this, participants may disengage, and the resultant system will miss out on valuable insights.

A variety of methods can be used to capture the requirements for the robot. Focus groups are often used as they are expected to generate a broader discussion; however the literature does not necessarily support this, but does indicate that a difference in the type of disclosures participants make is present [20]. One-to-one interviews allow a very in-depth insight into a participants' thoughts, but the interviewer can often influence the interview with their own biases [21]. Various workshop techniques such as Art of the Possible or 635 brainwriting [22] leverage divergent thinking to generate a wide range of ideas in a short time period. Observation is another method as it allows a roboticist to actively see the flow of a customer interaction, from the minute a user walks in to the door to their interaction with an employee ending. A mixed-methods approach is likely an ideal scenario.

Another risk is that stakeholders do not understand the current state of robotics and the implications that may arise. They need to be informed that the robot will fail, and assured when it does fail, that it is fail-safe and fail-closed [23]. To consider any robot not prone to failure is dishonest and untrue. For example, a recent deployment in a hat shop had a failure rate of 30% [24]. Not only does this cause the possibly for customer dissatisfication, it can also have legal ramifications. In the case of Moffat vs Air Canada [25], the airline were held liable for the responses of their chatbot despite the chatbot not being explicitly programmed incorrectly; the LLM simply hallucinated its response. This becomes a two-fold problem in the real-world; not only is language still an issue, but there is also risk of the robot causing physical harm. Therefore, just as in other software projects, the stakeholders do not need to understand the technical components innately, but they must be aware of the limitations and concerns of a design decision; facial analysis being one such example of this.

The definitive stakeholders should provide their own acceptable level of risk as a way to shape the design of the system. If, for example, it is acceptable that the robot may give erroneous information, such as a robot there to assist with chit-chat at the entrance to a shop, high temperature LLMs could be used to provide a more engaging dialogue system. However, if the stakeholder says no mistakes are allowed, such as in the case of a robot taking card payments, a more conservative or even rule-based dialogue system could be required at the expense of the robot freely conversing.

Stakeholders should also be asked about what they consider as their measure of success for the robot. The notions of value between different stakeholders will likely vary, and understanding this at the outset allows us to truly assess the robot's performance. For example, management may desire increased visitor numbers when a robot is deployed outside a shop as their metric, whereas end users want their wait times for service decreased, and front-line workers want to stop answering repetitive queries. As researchers, we often simply rate performance with end users against classical interaction scales, and possibly a custom survey to answer a specific research question. To truly begin to understand how to integrate robots into real-world spaces, we have to get feedback from all of our stakeholders. This allows us to begin to understand not only what different groups value in a robot, but learn how to manage these competing priorities in our system design.

4. Why?

Human-centered design is important as not only does it solve business objectives, but also empowers humanity at large. Industry 5.0, a phase of industrialisation, is partly defined by the European Commission as "adopting a human-centric approach for digital technologies including artificial intelligence" [26]. From a roboticist's point of view, HCD promotes them to consider ethical and sustainable design, as they have to acknowledge the impact of their robot within a space. It also helps to reduce barrier to entry, as the design has been informed by people that understand the deployment context. Public spaces are one of the most diverse spaces in our human world [27], so creating a system that can perform in this environment requires us to consult a diverse range of people. Our goal is to provide a user experience that users not only enjoy, but can also extract the information they require and thus complete the task, whilst making it possible for anyone to do. Therefore, for truly human-centered design of public-space robots, we have to involve and consider all stakeholders in the design process.

For example, if front-line workers are not consulted, it is likely that the true essence of the customer-worker interaction is lost. In wider service design projects, Karlsson found that front-line workers contribute three distinct types of knowledge: customer, product and practice [28]. Customer knowledge in the physical domain can be considered akin to logging and usability metrics within the digital domain; it allows us to see reoccurring problems users experience, and opportunities not only to fix these problems but also to improve the overall service.

However, if workers are not consulted in the right way, there exists the risk of "knowledge-hiding". This is where the employees will purposely refrain from sharing experiences. This can be seen between colleagues when competing for promotion [29]. Worryingly, this is also seen when job security is threatened, which is a key and recurring issue in robotic deployment. To address this, it is important to clarify to all stakeholders that robot deployments do not seek to replace workers; they aim to take on the repetitive tasks to free them to spend more time on more meaningful and complex tasks. This must be communicated through good leadership, however evidence for moderating leadership tactics being able to mitigate this is unclear and requires further work [30]. In practice, this factor is often mitigated when workers actually interact with the robot; they often even become fond of the robot, such as a team of soldiers who gave a funeral to a destroyed military robot [31].

This wide net of stakeholder engagement allows us to stop our robotic deployments being "red-carded". This term, akin to that in sports, reminds us that at any point a robot deployment can be stopped if the robot has been deemed to commit a serious violation that threatens a stakeholders reputation or wellbeing. If the robot has zero interactions with end users, then it will likely be stopped as it is a waste of resource for an organisation, and possibly even embarrassing if it cannot perform. If it offends end users, such as with poorly designed humour, it will likely be stopped to prevent - valid - complaints. A robot is not seen as a free entity like a human staff member; it is simply another piece of technology and should be held to the same standards as any other. We need to design a system that can operate autonomously and perform its given task; whilst also satisfying all stakeholders simultaneously. We note this is no simple task; but it is a near impossible task if we do not consider all of our stakeholders in the first instance.

5. Conclusion

As robots continue to move from the laboratory to the real world, we propose that the key to unlocking the potential of autonomous social robots lies in wide stakeholder consideration and engagement. To achieve this, researchers and industry need to identify a much wider range of stakeholders and prioritise them accordingly using established stakeholder theory, design and perform thorough stakeholder engagement methods to elicit domain knowledge, and therefore allow feasible and meaningful tasks for the robot to be selected by the stakeholders themselves. With this, we believe success of robotic deployments will increase and allow us to increasingly learn and begin to reliably address real-world concerns with robot deployments. We look to apply this in our own future work, and provide HRI-specific stakeholder and design theory guidelines for the field.

References

[1] M. E. Foster, B. Craenen, A. Deshmukh, O. Lemon, E. Bastianelli, C. Dondrup, I. Papaioannou, A. Vanzo, J.-M. Odobez, O. Canévet, Y. Cao, W. He, A. Martínez-González, P. Motlicek, R. Siegfried, R. Alami, K. Belhassein, G. Buisan, A. Clodic, A. Mayima, Y. Sallami, G. Sarthou, P.-T. Singamaneni, J. Waldhart, A. Mazel, M. Caniot, M. Niemelä, P. Heikkilä, H. Lammi, A. Tammela, Mummer: Socially intelligent human-robot interaction in public spaces, in: Proceedings of AI-HRI 2019, Arlington, VA, 2019. URL: https://arxiv.org/abs/1909.06749.

[2] A. Andriella, C. Torras, G. Alenyà, Short-Term Human–Robot Interaction Adaptability in Real-World Environments, International Journal of Social Robotics 12 (2020) 639–657. URL: https://doi.org/10.1007/s12369-019-00606-y. doi:10.1007/s12369-019-00606-y.

[3] S. Thunberg, T. Ziemke, Are people ready for social robots in public spaces?, in: Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, 2020, p. 482–484. doi:10.1145/3371382.3378294.

[4] L. De Gauquier, M. Brengman, K. Willems, H.-L. Cao, B. Vanderborght, In or out? A field observational study on the placement of entertaining robots in retailing, International Journal of Retail & Distribution Management 49 (2021) 846–874. URL: https://doi.org/10.1108/IJRDM-10-2020-0413. doi:10.1108/IJRDM-10-2020-0413, publisher: Emerald Publishing Limited.

[5] J. Wright, Robots Won't Save Japan: An Ethnography of Eldercare Automation, Cornell University Press, 2023. doi:https://doi.org/10.7591/cornell/9781501768040.001.0001.

[6] F. Nijhof, 2024.6: Dipping our toes in the world of AI using LLMs, 2024. URL: https://www.home-assistant.io/blog/2024/06/05/release-20246/.

[7] F. Zhang, F. Broz, E. Dertien, N. Kousi, J. A. M. van Gurp, O. I. Ferrari, I. Malagon, E. I. Barakova, Understanding Design Preferences for Robots for Pain Management: A Co-Design Study, in: Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction, HRI '22, IEEE Press, Sapporo, Hokkaido, Japan, 2022, pp. 1124–1129.

[8] E. Malnatsky, S. Wang, K. V. Hindriks, M. E. Ligthart, Shaping Relatable Robots: A Child-Centered Approach to Social Personalization, in: Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction, HRI '24, Association for Computing Machinery, New York, NY, USA, 2024, pp. 127–129. URL: https://doi.org/10.1145/3610978.3638374. doi:10.1145/3610978.3638374.

[9] R. Freeman, J. Mcvea, A Stakeholder Approach to Strategic Management, SSRN Electronic Journal (2001). doi:10.2139/ssrn.263511.

[10] J. Leventon, L. Fleskens, H. Claringbould, G. Schwilch, R. Hessel, An applied methodology for stakeholder identification in transdisciplinary research, Sustainability science 11 (2016) 763–775.

[11] H. Sharp, A. Finkelstein, G. Galal, Stakeholder identification in the requirements engineering process, in: Proceedings. Tenth International Workshop on Database and Expert Systems Applications. DEXA 99, 1999, pp. 387–391. doi:10.1109/DEXA.1999.795198.

[12] R. K. Mitchell, B. R. Agle, D. J. Wood, Toward a Theory of Stakeholder Identification and Salience: Defining the Principle of Who and What Really Counts, The Academy of Management Review 22 (1997) 853–886. URL: https://www.jstor.org/stable/259247. doi:10.2307/259247, publisher: Academy of Management.

[13] A. Mendelow, Stakeholder mapping, in: Proceedings of the 2nd international conference on information systems, A. Mendelow Cambridge, MA, 1991, pp. 10–24.

[14] C. Zaga, M. L. Lupetti, D. Forster, D. Murray-Rust, M. Prendergast, D. Abbink, First International Workshop on Worker-Robot Relationships: Exploring Transdisciplinarity for the Future of Work with Robots, in: Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction, HRI '24, Association for Computing Machinery, New York, NY, USA, 2024, pp. 1367–1369. URL: https://dl.acm.org/doi/10.1145/3610978.3638156. doi:10.1145/3610978.3638156.

[15] M. E. Foster, S. Ashkenazi, A. Blair, A. Ramirez-Duque, S. Ali, P. Candelaria, S. Hudson, J. Stinson, L. Harris, S. Litwin, F. Nishat, R. P. A. Petrick, A. Lindsay, D. Harris Smith, A. Dumlu, F. Zeller, Including Front-Line Workers as Primary Stakeholders in Public-Space HRI, Boulder, CO, USA, 2024. URL: https://eprints.gla.ac.uk/321922/.

[16] D. Brščić, H. Kidokoro, Y. Suehiro, T. Kanda, Escaping from children's abuse of social robots, in: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI '15, 2015, p. 59–66. doi:10.1145/2696454.2696468.

[17] P. Serrador, R. Turner, What is Enough Planning? Results From a Global Quantitative Study, IEEE Transactions on Engineering Management 62 (2015) 462–474. URL: https://ieeexplore.ieee.org/abstract/document/7165629. doi:10.1109/TEM.2015.2448059, conference Name: IEEE Transactions on Engineering Management.

[18] C. Parkinson, Parkinson's Law, and Other Studies in Administration, Buccaneer Books, 1957. URL: https://books.google.co.uk/books?id=SwbYAAAAMAAJ.

[19] J. Cao, How Robot Comes into Our Life in Urban Public Space: A Participatory Study, in: Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction, HRI '24, Association for Computing Machinery, New York, NY, USA, 2024, pp. 292–296. URL: https://dl.acm.org/doi/10.1145/3610978.3640638. doi:10.1145/3610978.3640638.

[20] G. Guest, E. Namey, J. Taylor, N. Eley, K. McKenna, Comparing focus groups and individual interviews: findings from a randomized study, International Journal of Social Research Methodology 20 (2017) 693–708. URL: https://doi.org/10.1080/13645579.2017.1281601. doi:10.1080/13645579.2017.1281601, publisher: Routledge _eprint: https://doi.org/10.1080/13645579.2017.1281601.

[21] P. T. Yanos, K. Hopper, On 'False, Collusive Objectification': Becoming Attuned to Self-Censorship, Performance and Interviewer Biases in Qualitative Interviewing, International Journal of Social Research Methodology 11 (2008) 229–237. URL: https://www.tandfonline.com/doi/abs/10.1080/13645570701605756.

doi:10.1080/13645570701605756, publisher: Routledge.

[22] B. Rohrbach, Kreativ nach regen-methode 635, eine neue technik zum losung von problemen, Absatzwirtschaft 12 (1969) 73–75.

[23] D. Liu, C. Li, J. Zhang, W. Huang, Robot service failure and recovery: Literature review and future directions, International Journal of Advanced Robotic Systems 20 (2023) 17298806231191606. URL: https://doi.org/10.1177/17298806231191606. doi:10.1177/17298806231191606, publisher: SAGE Publications.

[24] S. Edirisinghe, S. Satake, D. Brscic, Y. Liu, T. Kanda, Field Trial of an Autonomous Shopworker Robot that Aims to Provide Friendly Encouragement and Exert Social Pressure, in: Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction, HRI '24, Association for Computing Machinery, New York, NY, USA, 2024, pp. 194–202. URL: https://dl.acm.org/doi/10.1145/3610977.3635007. doi:10.1145/3610977.3635007.

[25] Moffatt v. Air Canada, 2024. URL: https://canlii.ca/t/k2spq.

[26] Industry 5.0 - European Commission, 2022. URL: https://research-and-innovation.ec.europa.eu/research-area/industrial-research-and-innovation/industry-50_en.

[27] M. Mintrom, S. Sumartojo, D. Kulić, L. Tian, P. Carreno-Medrano, A. Allen, Robots in public spaces: implications for policy design, Policy Design and Practice (2022). URL: https://www.tandfonline.com/doi/abs/10.1080/25741292.2021.1905342, publisher: Routledge.

[28] J. Karlsson, P. Skålén, Exploring front-line employee contributions to service innovation, European Journal of Marketing 49 (2015) 1346–1365. URL: https://doi.org/10.1108/EJM-10-2012-0568. doi:10.1108/EJM-10-2012-0568, publisher: Emerald Group Publishing Limited.

[29] T.-M. Nguyen, A. Malik, P. Budhwar, Knowledge hiding in organizational crisis: The moderating role of leadership, Journal of Business Research 139 (2022) 161–172. URL: https://www.sciencedirect.com/science/article/pii/S0148296321006718. doi:10.1016/j.jbusres.2021.09.026.

[30] Y. Xie, Q. Xia, J. Song, S. Hu, X. Liu, How ethical leadership influences knowledge hiding? A sequential mediation model, The Service Industries Journal (2024). URL: https://www.tandfonline.com/doi/full/10.1080/02642069.2023.2245356, publisher: Routledge.

[31] J. Smids, S. Nyholm, H. Berkers, Robots in the Workplace: a Threat to—or Opportunity for—Meaningful Work?, Philosophy & Technology 33 (2020) 503–522. URL: https://doi.org/10.1007/s13347-019-00377-4. doi:10.1007/s13347-019-00377-4.