What we Learn on the Streets: Situated Human-robot Interactions from an Industry Perspective

Anna Dobrosovestnova
Human-Computer Interaction Group, TU Wien
Vienna, Austria
anna.dobrosovestnova@tuwien.ac.at

Astrid Weiss
Human-Computer Interaction Group, TU Wien
Vienna, Austria
astrid.weiss@tuwien.ac.at

ABSTRACT

As commercial robots are increasingly present in everyday spaces, more efforts in the HRI community are directed towards investigating the complexity of interactions in naturalistic settings. Simultaneously, conversations about the potentials and challenges of collaborations between academic HRI and industry are taking place. This paper contributes to these topics by presenting two – *'People not Users'*, and *'What we Learn on the Streets'* – out of the five themes qualitatively developed from the interviews conducted with a company developing and deploying commercial delivery robots in public spaces.

CCS CONCEPTS

• **Human-centered computing → HCI design and evaluation methods**; **HCI theory, concepts and models**;

KEYWORDS

industry, expert interviews, developing robots for the real world, beyond users, industry and academia relationships

ACM Reference Format:
Anna Dobrosovestnova and Astrid Weiss. 2024. What we Learn on the Streets: Situated Human-robot Interactions from an Industry Perspective. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24 Companion), March 11–14, 2024, Boulder, CO, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3610978.3640753

1 INTRODUCTION

The increasing presence of commercial robots in public spaces has prompted studies in the Human-Robot Interaction (HRI) community to explore real-world interactions (see e.g., [9, 10, 13]). These studies highlight the complexity of the situated HRI compared to controlled laboratory settings. Simultaneously, efforts to understand commercial robot development and its distinctions from academic HRI goals are underway (e.g., [12], [6]). This report contributes to these efforts by examining insights from three interviews with representatives of Starship, a company developing autonomous delivery robots, conducted as part of the first author's PhD research project exploring situated interactions with commercial technologies in public spaces. The interviews covered topics such as robot design evolution, challenges faced when developing a commercial robot for deployment in public spaces, acceptance, robot sociality, and others. The qualitative analysis of the data yielded five themes, with two – *'People not Users'*, *'What we Learn on the Streets'* – presented here. The paper concludes with key learning points and open questions inspired by the data analysis and its outcomes.

HRI '24 Companion, March 11–14, 2024, Boulder, CO, USA

ACM ISBN 979-8-4007-0323-2/24/03.
https://doi.org/10.1145/3610978.3640753

2 PARTICIPANT RECRUITMENT, DATA COLLECTION AND ANALYSIS

After unsuccessful attempts to establish contact with the company through publicly available contact channels, a shared contact facilitated an introduction to a key company figure. This introduction allowed the first author to visit the company and explain the research project in person. The initial meeting involved three individuals closely involved in robot development (one of the company's co-founders, and representatives of the navigation and autonomous driving teams). Subsequent separate interviews with each of these three people took place at the company's office, lasting approximately one hour each and recorded for analysis. At the time, a formal consent form was not signed, as the interviewees requested us to obtain their agreement on publication content before release. Despite these deviations from standard procedures in academic HRI, all interviewees actively participated in the discussions and were open to sharing their experiences.

For the analysis of the interviews, the first author was guided by Reflexive Thematic Analysis (RTA) as developed by Virginia Braun and Victoria Clarke [2]. Situated within a qualitative research paradigm, compared to other approaches to thematic analysis such as e.g., coding reliability approaches that lean towards more quantitative and positivist paradigms [1, 3], the focus in RTA is on *rich meaning* and generation of contextualised and situated knowledge [2, p.6]. The coding process in RTA is organic, with a potential for the codes to evolve with the researcher's deepening understanding of the data. Subsequently, coding and theme development in RTA rely on the interpretative work of researcher [1]. The first author chose RTA because it aligned with the critical constructivist epistemology within which she situates her research, and allowed flexible exploration of the data through entangling descriptive and interpretative approaches. A mixture of deductive and inductive coding was used, with the first iteration of codes developed while reading through the printed interview transcripts. The subsequent iterations of coding were done in the MAXQDA software (Version 2022.0.0). Based on the resulting codes, five broader themes were iteratively developed. The remainder of the report focuses on two of these: *'People not Users'*, and *'What we Learn on the Streets'* as these themes offer insights particularly relevant in the context of the broader conversation about robots in the real world [1].

3 REPORTING OF THE THEMES

3.1 People not users

At the core of this theme was a shared contention by the interviewees that 'user' as a term falls short against the plurality of actors who come to experience and interact with the robots in different ways and to different effects for the ongoing robot development process. Instead of talking about 'users', the interviewees differentiated between two broad groups of actors – the company's customers, and people more broadly: *"We try not to use the word. We have customers, and then otherwise, we just talk about people. Because [it] doesn't seem like the right word."* (P2). Within this distinction, customers were identified as people who have some sort of *formal* relationship with the company. Customers use the company's app to place orders with vendors who use the robots as a delivery option [2]. Given the nature of their involvement as customers, features of the robot and supporting infrastructure that are commonly considered the domain of User Experience (UX) are especially relevant when it comes to this group of actors.

When it comes to the group delineated broadly as *people*, the reasoning behind the choice of our interviewees to refrain from calling them users had to do with the perceived inadequacy of the term to capture the nature of their very heterogeneous relationships to and interactions with the robots, which exceeded the space conventionally demarcated as "use": *"Someone who passes by on the street, or a driver, they aren't using [the robot] in the same way that they aren't using anyone else they passed by on the street, or they're not using the pedestrians they stopped in front of at a pedestrian crossing."* (P2). Despite the lack of a formal relationship between the company and this group, *people* are of significant importance in the robot development process. People are the ones populating spaces in which the robots have to integrate: *"And so you have just one percent of interactions that [we] know anything about, the formal interaction [...] Right now, every other time, it's someone with some thought about what the robot is, and some intent towards it, some mental state, and we can't say much more about it unless we figure it out. And it's a fascinating problem because it means that most of the work that we do on this affects people without relationships to us and without feedback to us. Without a way to express their satisfaction with it, or any reason to do so."* (P2). Despite a certain passivity suggested by the statement above, people can and do engage in proactive interactions that can hinder or contribute to the robots' operations (more on this in 3.2). Furthermore, people form local communities that either accept or reject the robots: *"And do we consider them? We consider them as part of the environment we are in, more than just part of the environment, part of the community. And robots have to successfully integrate into this community of people. Otherwise, we have not reached our goal as a business."* (P1).

In some ways invisible and inscrutable from the perspective of the company, people are at the same time not a homogeneous group. Among the many different sub-groups constituting *people*, passersby, drivers, cyclists, elderly people, children, and people with disabilities were explicitly mentioned. The differentiation between these groups of actors stemmed from the interviewee's experiences that people interact with the robots on different terms even when interactions are reduced to being 'incidentally co-present' [11] in the environment. For instance, as explained by P1, drivers, cyclists and pedestrians rely on different stimuli and modalities to make their way through the environment: *"We've seen, for example, robots commonly cross the road. And [...] we have [...] distinct categories of people who interact with [the robot]: pedestrians and sidewalk users, and cyclists. And then drivers. And drivers are special because they're in cars and they are in a structured road environment and they follow rules; whilst pedestrians, they just walk in, there are no real rules to a sidewalk environment. [...] So a cyclist might not see the robot because it is hidden, but they are actually very attuned to hearing something, that oh, there is actually something around, there's maybe an E-scooter or something like this. Cyclists very much rely on hearing cars and they go by this".* The differences and certain degree of unpredictability to how these different groups of actors navigate the space and interact with the robots pose a challenge for decision-making about robot features and behaviors e.g., how to ensure the robot is predictable, understandable and, ultimately, *safe* for everyone involved.

Which sub-groups come to matter more, in which manner, and why, also differed. Elderly people, who were considered more vulnerable compared to an average adult person, were emphasized as one sub-group who were considered extensively in the design process: *"Generally, they are a little bit less agile on the sidewalk or on the road. So there might be more of a concern from our side [what] we actually want to behave like. Keep enough distance from them, don't swerve in their path, and so on... If an elderly person falls because of a robot – this is a scenario which we've considered thoroughly: how dangerous is the scenario and how often would [it] happen? How can we prevent this from happening effectively? So far, there hasn't been any major incident that I know of. But this is for sure something we consider."* (P1). Disabled people were another group warranting special consideration. Extending considerations to this group of actors meant proactively engaging with communities of people with disabilities to collaboratively work out the best strategies for robot navigation: *"I would say people in wheelchairs, or otherwise disabled people, have been much more vocal as a group. And we have engaged with them in most places that we go, in order to try and figure out how best to behave. Because in this case, both the robot and a person in a wheelchair are somewhat limited as to what paths they can take. Where a person on their feet is less limited – they can step off the sidewalk much more easily, or step into the road or onto a curb. But a robot can't easily go off a curb and back onto it. And obviously, neither can a person in a wheelchair. And so I think there is much more engagement and trying to figure out how best to cooperate. And we have initiatives ongoing right now to improve the detection of mobility devices, strollers and wheelchairs."* (P3).

While these two groups of actors receive special considerations due to the perceived vulnerability they share, the example of people with disabilities also points to the role that the power to consolidate

[1]The three themes outside the scope of this report are: *'Robot Development and Design as an Ongoing Process of Negotiations'*, *'Robots are a different kind'*, and *'Behind the scenes of automation'*

[2]Note that vendors are also a group who could be considered customers, but they were not explicitly discussed in the interviews with the company. We are, however, preparing a separate publication that includes the company's partners' perspective.

and voice opinion, potentially contributing to the overall robot acceptance or rejection, plays in shaping who gets included in decision-making about robot design and behavior, and how. To us, this suggested that some groups of implicated actors (e.g., human delivery couriers) – though affected by robot deployment – may remain only partially visible or considered due to their different social and political positioning in the context of a local community.

3.2 What we learn on the streets

Throughout the interviews, the interviewees emphasized the differences, as they perceived and experienced them, between academic HRI and robot development in an industry setting. Among these differences, requirements for scalability, robustness (including safety), and overall financial goals and constraints played a critical role in shaping the overall trajectory of the robot development: *"[..] If you put in restrictions on the size, the weight, the cost to produce, reliability, and some other things [...] [This] is just very different between an academic setting and a high volume commercial setting, which needs to be [at] a totally different level in terms of how easy it is to do and how reliable it is – and how low cost it is."* (P3). These constraints, coupled with the fact that the robots are intended for successful operations in unconstrained environments, render much of the knowledge generated within academic HRI pragmatically irrelevant. As shared by P1, even though they try to stay in touch with the academic community and do rely on HRI studies up to a certain point, when it comes to solving challenges in the context of robot deployment in the real world, few academic studies apply: *"But we do then try to base it on studies and on literature. And we have reviewed that and I try to keep up with this literature as well. But there is [so] much of it. And not much of it actually applies to what we do."*

The perception of the limited transferability of the academic knowledge was related, apart from business considerations, to a significant reduction of complexity characterizing much of the HRI research. Above, we already considered the example of ensuring that the robot is predictable and understandable for a plurality of dynamic actors in the environment. Even when one succeeds to identify a behavior strategy that is a good enough fit, how people will respond to it in a situated manner can never be fully predicted or anticipated. For example, when the robots were first introduced, a common situation emerged where drivers and the robot were both waiting at crossroads to let the other pass. With time, the drivers learned that, even though existing regulations may obligate them to let the robot pass [3], it makes no pragmatic sense to do so as the robots will wait too (because they are programmed to err on the side of caution).

Another potential impediment to transferability had to do with the fact that what works in an idealised lab setting may not integrate well into real world settings. One example shared by P1 concerned intent signaling: *"So we know that, or we expect, from various studies, that the strongest indicator a robot has to communicate its own intent is movement. But movement actually requires moving, and moving in some situations can be dangerous or is highly controlled – we don't want the robot to move uncontrollably, we don't want the robot to speed up, we don't want the robot to hit things. So if you come up with the greatest ever pattern to show that the robot wants to move left, then there is still a constraint that actually, first and foremost, the robot has to be safe and has to remain autonomous. And only then, if it has enough space, it could actually show some pattern that it wants to move left or something like this. That's why we have often used lesser methods such as lights to indicate what the robot might do, or wants to do. Because they're very safe and easy. They cannot break anything by flashing lights."*.

In 3.1, we touched upon how people as a heterogeneous group come to matter, not in the least because of the various forms of proactive engagements with the robots. As pointed out by the interviewees, much of what transpires between people and robots on the street was surprising, or could not have been fully anticipated in advance: *"I can rationalize it for myself after the fact. But I could not have predicted it. I would say I'm not shocked by it. I'm not surprised, after the fact, I think it actually makes sense, I can explain it for myself why this happens. But you know, at least back then, without having actually experienced or seen this in the world, I would not have guessed that this will happen."* (P3 about the instances of people assisting the robots) [4].

A particularly interesting facet to explore in the data was how decisions were made in terms of altering the robot's features and behaviors in response to patterns of situated interactions on the streets. Deciding whether to act [5], and how to act, involved some form of (ideally, quantified) estimation of the overall frequency of occurrences, the impact the behaviors had, and the efforts – including skills, time and financial resources – required to implement the change (provided it is possible to identify what this change could be (which is not always the case) compared with the potential benefit of implementing a change. For example, LED flags being broken was one instance towards which systematic efforts were directed in the process of iterating on the robot design: *' "People can also touch our flag without the intent, malicious intent, right? So I would say that it's not necessarily super obvious what we should do? Apart from making the flag cheaper, which we are going to do actually. But at least it's a clear, systematic, and actually quantifiable, you know, thing to work on. Significant enough to very systematically work on."* (P3).

Whereas the company's interest in mitigating destructive behaviors towards robots seems straightforward, considering assistive behaviors as well as generally positively charged social encounters with the robots in the design process, or deciding whether to attempt to elicit more of these behaviours, was a subject of ongoing negotiations. A pertinent example concerned people voluntarily helping the robots – an instance of behavior that was first observed (to the interviewees' shared surprise) and only then integrated into the robot's communication strategies – the robots can now proactively solicit the help of passersby through dialogue-based strategy. Furthermore, tensions emerged in the way different interviewees reasoned what the help of passersby meant for the company. On the one hand, the interviewees acknowledged that people assisting the robots was something that was "nice to have", and in some cases it could even be useful: *"This project to ask people to press the button*

[3]In Estonia, where the study was conducted, the robots were legally equated to pedestrians.

[4]For more on the topic of passersby assisting Starship robots see [4]

[5]In this case, acting assumes addressing systematic efforts as well as resources towards a specific problem.

at the crossing will be like a pioneering ask, but this is not going to be... it's gonna be a relaxed–I would say, it's going to be nice to have, it's going to speed things up." (P2). On the other hand, when asked whether they rely on such behaviors, the answer was rather negative: *"But that's a very small minority of the times. And that's days in a year [...]. So the answer is: not a great deal."* (P2). The reasons for this were twofold: firstly, as suggested by the statement above, the overall number of cases when the robots require external assistance are rather low compared to the overall distance driven every day. Secondly, including seeking the help of passersby, though not difficult to implement technically speaking, was in certain tension with the company's overall strategy directed at minimizing dependency where possible. In other words, as explained by P2, as a rule, even when a shortcut is possible (e.g., by having the robot ask a human to press a traffic light every time it needs to cross a street), the default choice will be to take a longer route: *"And the same way, if you can't rely on these things, if you can't rely on people always being there at all, let alone wanting to be helpful about it. So we just build the system to not need it. And then in general, this means – for example, it means we intentionally route robots around, we will wrap robots around crossings that need the button to be pushed. We'd rather go a long way, instead of going to a crossing that would be, you know, five minutes faster."* (P2).

The fact that the option of the robot asking for help was nevertheless integrated was framed by an interviewee more as an afterthought, rather than something that stemmed from serious considerations and substantial evaluations of impact (as was the case with the breaking of the flag): *"It was afterwards. We did not think that, hey, maybe we can induce people to help robot this way. If the robot asks for help. But no, I will say this has been very much an afterthought, or very much something where we have reacted to what we see in the world. And if anything, I would say we have reacted very slowly, in this area.'* (P2). Importantly, what makes such "afterthoughts" and side projects possible is that they are appraised as rather low cost, easy to integrate on the existing software and hardware pipeline, and reversible if proved useless. They also do not interfere in any substantial way with fundamental concerns such as autonomous navigation, reliability and safety. In other words, even if not perceived as high on the priority list, efforts directed at enabling different forms of positively charged quasi-social interactions, though less systematic, still take place, not in the least because of the recognition that such interactions – despite challenges in quantifying their impact – contribute positively to the overall perception of the robots as new actors in the community: *"So it's a nice add-on from a utilitarian perspective. But not the game changer. I think it actually has a bigger impact in terms of how the robots are perceived ... We're also seeing, you know, firsthand, that people like that the robots are interacting with them in more ways than just moving or stopping or treating them as an obstacle, but actually, interacting in terms of actually saying something. Okay, it's one-way communication: they're responding, but the robot doesn't really hear. But even that sort of, very rudimentary sort of interaction or communication, we certainly see that people just like it a lot. And it's very cheap for us to have it."* (P3). The underlying constant in how the interviewees reasoned about decisions related to the development of social dimension of the HRI remained the contention that, should any form of systematic efforts be undertaken to enable more social interactions with the robots, these will have to be coupled with the same quantification-oriented approach characterizing the ongoing developments that are rated high on the priorities list: *"In future, we may actually have the time to have these interactions, we may understand the trade-offs of them more. You could foresee things like that: we could actually understand what times of day there are enough people there to press the button. [...] And then we can build this behavior. So the effect is not in, you know, not sending in the morning, when there's no one around, and not right in the middle of the day, when everyone's pressing it every second anyway. But those times of the day [when it makes sense], then we could build this interaction in."* (P2).

4 DISCUSSION

In conclusion, our analysis has shed light on the distributed nature of a commercial robot development process involving many diverse stakeholders and blurring the divide between the development and interaction space, wherein the latter integrates but is not limited to what we commonly demarcate as *use*. Concerning customers as related stakeholders, we ague that the field of HRI lacks a comprehensive framework for customer experience (CX) research, with an existing emphasis primarily on understanding UX [8]. Our first proposal for future emphasis in academia is to develop a robust framework that encapsulates the multifaceted dimensions of CX in the realm of robotics. Concerning broader groups of actors beyond customers, while acknowledging predecessors such as the EU project REELER, which presented the Human Proximity Model derived from ethnographic research where various stakeholders were considered [5], we point towards another critical gap in the current academic HRI landscape. Namely, by framing related actors as either users or merely incidentally co-present, we risk impeding opportunities for more nuanced understanding of how different actors come to matter to everyday robot operation and iterations on design. This revelation exposes another noteworthy gap in academic discourse: the iterative nature of robot development [15]. Aside from few examples, such as Snackbot [7] and IURO [14], which prioritized and reported iterative design for robotic solutions, we observe a scarcity of studies dedicated to focusing on the wicked nature of HRI problems – when you solve one problem in the design, the next one might pop right up. Our call for attention to the wicked nature of developing robots underscores the value for the academic community in appreciating the nuanced complexities of each project. While we do not dismiss the relevance of laboratory studies, we posit that for industry uptake, more HRI studies must prioritize real-world constraints, addressing factors like lighting conditions, ambient noise, and dynamically changing environments. On a related note, as emphasized by the example of limited functional value of passersby assisting robots on the streets as perceived by developers, this also means an ongoing evaluation of the actual importance for the real-world deployment of certain trajectories of research pursued within academic HRI.

ACKNOWLEDGMENTS

The first author is a recipient of a DOC Fellowship of the Austrian Academy of Sciences (DOC/25861)

REFERENCES

[1] Virginia Braun and Victoria Clarke. 2021. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research* 21, 1 (2021), 37–47. https://doi.org/10.1002/capr.12360 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/capr.12360.

[2] Virginia Braun and Victoria Clarke. 2021. *Thematic Analysis : A Practical Guide.* 1–100 pages. Publisher: SAGE Publications Ltd.

[3] Virginia Braun and Victoria Clarke. 2023. Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a *knowing* researcher. *International Journal of Transgender Health* 24, 1 (Jan. 2023), 1–6. https://doi.org/10.1080/26895269.2022.2129597

[4] Anna Dobrosovestnova, Isabel Schwaninger, and Astrid Weiss. 2022. With a Little Help of Humans. An Exploratory Study of Delivery Robots Stuck in Snow. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN).* 1023–1029. https://doi.org/10.1109/RO-MAN53752.2022.9900588

[5] Cathrine Hasse, Stine Trentemøller, and Jessica Sorenson. 2018. The Use of Ethnography to Identify and Address Ethical, Legal, and Societal (ELS) Issues. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (Chicago, IL, USA) *(HRI '18).* Association for Computing Machinery, New York, NY, USA, 393–394. https://doi.org/10.1145/3173386.3173560

[6] Jin Joo Lee, Amin Atrash, Dylan F. Glas, and Hanxiao Fu. 2023. Developing autonomous behaviors for a consumer robot to be near people in the home. In *IEEE RO-MAN 2023.* https://www.amazon.science/publications/developing-autonomous-behaviors-for-a-consumer-robot-to-be-near-people-in-the-home

[7] Min Kyung Lee, Jodi Forlizzi, Paul E. Rybski, Frederick Crabbe, Wayne Chung, Josh Finkle, Eric Glaser, and Sara Kiesler. 2009. The Snackbot: Documenting the Design of a Robot for Long-Term Human-Robot Interaction. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction* (La Jolla, California, USA) *(HRI '09).* Association for Computing Machinery, New York, NY, USA, 7–14. https://doi.org/10.1145/1514095.1514100

[8] Jessica Lindblom, Beatrice Alenljung, and Erik Billing. 2020. Evaluating the user experience of human–robot interaction. *Human-Robot Interaction: Evaluation Methods and Their Standardization* (2020), 231–256.

[9] Frederik Moesgaard, Lasse Hulgaard, and Mads Bødker. 2022. Incidental Encounters with Robots. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN).* 377–384. https://doi.org/10.1109/RO-MAN53752.2022.9900591 ISSN: 1944-9437.

[10] Frederik Moesgaard, Lasse Hulgaard, and Mads Bødker. 2022. Incidental Encounters with Robots. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN).* 377–384. https://doi.org/10.1109/RO-MAN53752.2022.9900591

[11] Astrid Rosenthal-von der Pütten, David Sirkin, Anna Abrams, and Laura Platte. 2020. The Forgotten in HRI: Incidental Encounters with Robots in Public Spaces. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction* (Cambridge, United Kingdom) *(HRI '20).* Association for Computing Machinery, New York, NY, USA, 656–657. https://doi.org/10.1145/3371382.3374852

[12] A AAAI 2023 Spring Symposium. 2023. HRI in Academia and Industry:Bridging the Gap. https://sites.google.com/view/aaai-hri-bridge/home. [Online; accessed 1-December-2023].

[13] David Weinberg, Healy Dwyer, Sarah E. Fox, and Nikolas Martelaro. 2023. Sharing the Sidewalk: Observing Delivery Robot Interactions with Pedestrians during a Pilot in Pittsburgh, PA. *Multimodal Technologies and Interaction* 7, 5 (2023). https://doi.org/10.3390/mti7050053

[14] Astrid Weiss, Nicole Mirnig, Ulrike Bruckenberger, Ewald Strasser, Manfred Tscheligi, Barbara Kühnlenz, Dirk Wollherr, and Bartlomiej Stanczyk. 2015. The interactive urban robot: user-centered development and final field trial of a direction requesting robot. *Paladyn, Journal of Behavioral Robotics* 6, 1 (2015), 000010151520150005.

[15] Astrid Weiss and Katta Spiel. 2022. Robots beyond Science Fiction: mutual learning in human–robot interaction on the way to participatory approaches. *AI & society* 37, 2 (2022), 501–515.