

## Preface

This volume contains the Proceedings of *Robophilosophy 2022: Social Robots in Social Institutions*, the fifth event in the biennial *Robophilosophy Conference Series*, which was held in Helsinki, Finland, August 16–19, 2022. The conference series was initiated by an interdisciplinary research group at Aarhus University in August 2014, with an event that marked the beginning of the new field of robophilosophy. Robophilosophy is defined as “philosophy of, for, and by social robotics” – a new area of applied philosophy undertaken in close interdisciplinary contact with empirical research in HRI, technical design, and robotics engineering. After the second event in 2016, which also took place in Aarhus, the series began to travel internationally with periodic changes of conference locations. In 2018, the conference was held in Vienna, Austria. The 2020 conference was again organized by Aarhus University, but it was held fully online because of the COVID-19 pandemic. In 2020, at the early stage of planning the Helsinki conference, we expected that by 2022 the pandemic would have fully subsided, and we could return to the traditional in-person conference format. However, the pandemic lasted much longer than we expected. In late 2021, when the call for papers was already sent out, yet another wave of the pandemic hit, and we started to worry about not receiving enough submissions unless we allowed for online presentations as well. Hence, we found ourselves back at the drawing board, now saddled with the task of planning how to organize a hybrid conference instead of a purely on-site event. In February 2022, Russia invaded Ukraine. Any war is deeply shocking and reason to despair of humanity, but this war, after a period of peace of more than 75 years after the devastation of the Second World War, affected Europeans profoundly. It also raised the concern that potential participants may not feel comfortable travelling to a country sharing a long border with the attacker. However, while such extra-academic worries were constantly in the background, in our small world all went very well—we had around 200 participants from 29 countries and around 100 high quality research presentations. Best of all, most of the authors were willing to attend in person, and created a very concentrated, friendly and intellectually inspiring atmosphere. Even the fickle Finnish weather was cooperative and allowed for discussions in street cafés long into the night. Simply put, the conference was a success at all levels—and yet, we were always aware that it constituted a four-day bubble offering the delight of joint academic work, a temporary mental safe haven in the midst of the horrible war that is unfortunately still raging in Ukraine as we write this.

Each Robophilosophy Conference has a special theme and for the year 2022 we chose *Social Robots in Social Institutions*. The aim of social robotics is to create entities capable of social interaction with humans. This raises questions about the notion of sociality, because our standard notions of sociality presuppose that the participants of interactions are persons, not robots or other artificial agents. In human societies, many forms of social interaction have been institutionalised. Broadly speaking, institutions emerge from social practices that coordinate activities by establishing formal and informal rules which, in turn, state the goals and values they serve and assign roles and positions with corresponding rights and responsibilities. Institutions guide individuals to coordinate their actions and cooperate in a way that stabilizes the institutions and serves the goals of the society. One of the aims of the conference was to understand and to

critically evaluate how social robotics can be expected to transform, and partly already is transforming, institutional structures, institutional practices, and the institution–citizen interaction for instance in the fields of social and health care, education, science, media, and law.

After almost two decades of interdisciplinary research into social robotics and Human–Robot Interaction (HRI) we still lack a clear understanding and regulative directives for how to ensure that social robotics will contribute to a community’s resources for affording human well-being—to the practices and institutions in which members of a community experience justice, dignity, autonomy, privacy, security, authenticity, knowledge, freedom, beauty, friendship, sensitivity, empathy, compassion, creativity, and other socio-cultural core values, as these may be shared, or vary, across cultures. Central questions concerning the larger societal significance of social robots are more urgent than ever: How does the introduction of social robots into our social institutions shape them and their ability to perform their functions in the society? How should we understand sociality in the context of human-robot interactions? What about social roles and practices? If we use social robots in social institutions, what are the effects of trust in institutions, responsibility allocations, and questions concerning transparency? What are the ecological and environmental consequences of large-scale use of social robotics and its technological prerequisites such as smart devices, networks, and computing and memory resources? How to institutionally prevent potentially harmful consequences? How to implement understanding of human social institutions including their normative aspects into the development of social robots? *Robophilosophy 2022* advanced the discussion of these and related questions, in plenaries, session talks, workshops, and individual exchanges that only an in-person event can enable.

The *Robophilosophy Conference Series* aims at promoting interdisciplinary Humanities research in and on social robotics. So we were particularly pleased that researchers from other disciplines participated, especially also from the multidisciplinary areas of HRI studies and social robotics. These cross-overs from overlapping interdisciplinary areas are needed to combine all relevant perspectives and results and approach the complex task of finding pathways towards developing social robots in a responsible fashion. This, in fact, is the core message of the *Robophilosophy Conference Series*: Only if humanities and social science researchers join forces with the research community and practitioners in social robotics and HRI can we create futures worth living.

With the organization of the conference in Helsinki we aimed to contribute to this mission, by way of exploring and debating themes, topics, and questions related to “Social Robots in Social Institutions”, embracing both theoretical and practical angles, as is the peculiar character of all *Robophilosophy* conferences.

The articles and abstracts collected in these *Proceedings*, which comprise almost all the research contributions presented at the conference, investigate social robots in various institutional contexts including religious, legal, and care institutions. The conference was an invitation and an opportunity to philosophers and other researchers in humanities and social sciences, as well as researchers in social robotics and HRI, to explore together how interdisciplinary research can contribute to shaping a future where social robotics is guided by the goals of enhancing socio-cultural values rather than mere economic utilities.