

Roadmap for Responsible Robotics

Dagstuhl Seminar: 10th-15th September, 2023

Hello. We are looking forward to seeing you at Dagstuhl in September. As those of you who have attended these seminars previously will know, they are intentionally informal with participants from across a variety of disciplines and backgrounds. However, it is good to at least begin with some structure and direction, then see where the discussions take us and how the broad aims develop.

Theme: We are concerned with robots, across all their forms, particularly autonomous robots that can make their own decisions and take their own actions without direct human oversight. “Responsible Robotics” covering both the responsible development and deployment of robots and the responsibilities encapsulated within such autonomous robots, is appealing and important. If we are to see robots in our society then these notions will indeed be crucial. As there is no agreed route to achieving “Responsible Robotics” in practice our Seminar will involve discussions around this, what the issues are and what the barriers are, with the aim of providing a draft roadmap for “Responsible Robotics” by its end. We have initially structured the seminar around themes commonly associated with responsible robotics, namely *trust*, *fairness*, *predictability* and *understandability*, refining and extending these as necessary. Understanding the interaction between these elements is crucial to many advanced uses of autonomous robots especially when in close proximity to humans. However, the Seminar will not only provide a clarification and refinement of these issues but a roadmap to their implementation in real-world systems.

Structure: We plan, on the first day (Monday 11th) to have a range of introductory talks on our topics of trust, fairness, predictability and understandability. Specifically, two talks of around 40 minutes on each topic across the day. In addition, we will intersperse brief introductions of all participants throughout the day. IN PREPARATION FOR THIS PLEASE COULD EVERYONE (EXCEPT THE LONGER SPEAKERS) PREPARE A 3 MINUTE INTRODUCTION TO THEMSELVES AND THEIR WORK (3 SLIDES, MAXIMUM) TO BE GIVEN ON MONDAY.

Over the next day and a half (all day Tuesday and Wednesday AM), participants will be split into four groups corresponding to our four key topics. (N.B: Wednesday PM is reserved for a group activity.) These groups will meet separately with the aim of, for their particular topic:

1. describing the barriers for implementation;
2. distinguishing between those issues that are solved, those that will never be solved, and those that could potentially be solved;
3. drafting a roadmap of those issues that could be solved, together with the research breakthroughs needed to solve them.

On Thursday AM each of the four groups will present their draft roadmap and discussions about each will take place across all attendees. Then, on Thursday PM, we will discuss interactions between issues and will attempt to map out a combined and consistent roadmap from the partial roadmaps provided in the morning.

Finally, on Friday AM a discussion will be held to (a) identify priorities and (b) agree next steps.

Request: Please come with an open mind, ready for constructive discussions.

Organizers:

| | |
|----------------------|--------------------------------------|
| Anna Dobrosovestnova | (TU Wien, AT) |
| Michael Fisher | (University of Manchester, UK) |
| Nick Schuster | (Australian National University, AU) |
| Marija Slavkovik | (University of Bergen, NO) |

Unable to attend: Seth Lazar (Australian National University) and Astrid Weiss (TU Wien)

SCHEDULE

MONDAY:

09:00 Introduction
09:15 2 x 30 minute overview talks
10:15 5 x 3 minute introductions
10:30 BREAK
11:00 5 x 3 minute introductions
11:15 2 x 30 minute overview talks
12:15 LUNCH
14:00 FREE TIME
15:00 1 x 30 minute overview talk
15:30 5 x 3 minute introductions
15:45 BREAK
16:15 5 x 3 minute introductions
16:30 2 x 30 minute overview talks
17:30 Planning for Day 2 - splitting into groups
18:00 END

TUESDAY:

09:00 Recap of Plan - splitting into groups
09:15 Groups meet separately
10:30 BREAK
11:00 Groups meet separately
12:15 LUNCH
14:00 Progress so far? Issues? Suggestions?
14:20 Groups meet separately
15:15 BREAK
15:45 Groups meet separately
17:00 END

WEDNESDAY:

09:00 Progress so far? Issues? Suggestions?
09:15 Groups meet separately
10:30 BREAK
11:00 Groups meet separately
12:15 LUNCH
14:00 EXCURSION/ACTIVITY

THURSDAY:

09:00 Presentation: Trust
09:20 Presentation: Predictability
09:40 Presentation: Fairness
10:00 Presentation: Understandability
10:20 BREAK
11:00 Discussion - issues/overlaps/conflicts
12:15 LUNCH
14:00 Discussion - coalescing four "roadmaps" into one

FRIDAY:

09:00 Presentation: draft joint roadmap
09:20 Discussion - issues/priorities
10:20 BREAK
11:00 Next steps
12:15 LUNCH
14:00 CLOSE