Designing for Trust: Autonomous Animal - Centric Robotic & AI Systems

Alan Chamberlain
School of Computer Science,
University of Nottingham
Alan.Chamberlain@Nottingham.ac.uk

Steve Benford
School of Computer Science,
University of Nottingham
Steve.Benford@Nottingham.ac.uk

Joel Fischer
School of Computer Science,
University of Nottingham
Joel.Fischer@Nottingham.ac.uk

Pepita Barnard
School of Computer Science,
University of Nottingham
Pepita.Barnard@Nottingham.ac.uk

Chris Greenhalgh
School of Computer Science,
University of Nottingham
Chris.Greenhalgh@Nottingham.ac.uk

Ju Row Farr
Blast Theory
Ju@Blasttheory.co.uk

Nick Tandavanitj
Blast Theory
Nick@Blasttheory.co.uk

Matt Adams
Blast Theory
Matt@Blasttheory.co.uk

ABSTRACT

From cat feeders and cat flaps to robot toys, humans are deploying increasingly autonomous systems to look after their pets. In parallel, industry is developing the next generation of autonomous systems to look after humans in the home – most notably robot arms that might assist with all manner of domestic tasks. How might the animals and humans in these spaces engage with these and with each other? This research focuses upon the role that 'Trust' plays in autonomous animal-centric robotic systems and the ways in which we as researchers can further understand how to design, develop, and evaluate such systems by taking a Responsible Research and Innovation (RRI) approach. Understanding, designing for, and negotiating 'Trust' is complex, particularly in contexts where animals, humans, and intelligent systems (including Robotics, AI, & IoT) come together in a social context, in which bonds are created, friendships develop, and mutual care plays a part in the interaction and developing relationship between people and animals.

CCS CONCEPTS

• **Human Centered Computing**; • **Applied Computing**;

KEYWORDS

Autonomy, Animals, ACI Animal-Computer Interaction, Trust, Robotics, AI, HCI, RRI, Human-Computer Interaction, Interaction, Design, Responsible, IoT

1 INTRODUCTION

From feeders to cat flaps and robot toys, humans are deploying increasingly autonomous systems to look after their pets. In parallel, industry is developing the next generation of autonomous systems to look after humans in the home – most notably robot arms that might assist with all manner of domestic tasks relating to multi-species interaction. Van De Linden *et al.* [10] note that, "*increasingly entangled daily routines lead to close multi-species households where dog owners conceptualize technology as having a role to support them in giving care to their dogs*", it is the notion of *entanglement* and in particular multi-species entanglement that leads us to reflect upon the design of systems in such contexts and to push beyond human-centered approaches, and to coin a phrase think about multi-species-centered approaches to design. It is also worth noting that other research has examined the ways in which technologies such as social agents are interacted with by "*dog-owners and dogs*" [8], while other research has taken a multi-species-centered approach to understand the impact of social robots in the household [9].

As point of departure we reflect upon the role that 'Trust' plays in autonomous animal-centric robotic and AI systems, and the ways in which we as a researchers can further understand how to design, develop, and evaluate such systems by taking a multi-species approach to Responsible Research and Innovation (RRI) [20] and to expand upon other design practices, such as Speculative Design to explore inter-species communication and design, as has been explored in a recent workshop to reflect on the "Ethics and Power Dynamics in Playful Technology for Animals" [6]. There is much written on Trust, particularly in the Computer Supported Cooperative Work community (CSCW), *see* Clarke *et al*'s edited collection [3] for an introduction to the field. We feel that Trust is timely and key area to focus our attention on, due to the rapid technical development in the area of autonomous systems, and the evolving intertwined nature of activities in multi-species settings that include such autonomous technologies, such as AI and robotics,

where trust may be seen as a feature spanning and impacting upon the overall design, development and use of such systems.

1.1 Trust, Design & Responsibility

Trust is foundational, it shapes and forms the world that we inhabit, our beliefs, behavior and reality. Notions of 'trust' are non-static and can be highly context specific, knowing what models of trust-based interaction to apply to a given context is not a simple task [17]; while deploying autonomous technologies in "the wild" [2] [4] which are effective and useful beyond the confines of the research lab is difficult, due to the 'specifics' and changing nature of the setting and context, and yet it is the real-world settings that can truly inform the design of technologies which animals and people engage. The value of understandings informed by the real-world practices is in their ability to direct the overall design of a system, allowing people to develop a deeper understanding of the practices of both animals and humans in their natural settings, leading to a more responsible and grounded approach to design, that respects [16] animals, and their interactions in a day-to-day context. Understanding, designing for, and negotiating 'Trust' is complex, particularly in contexts where animals, humans, and intelligent systems (including Robotics, AI, & IoT) come together in a social context, in which bonds are created, friendships develop, and mutual care plays a part in the relationship between people and animals. Systems need to be reliable, trusted, and adaptable to work in settings that are 'critical' such as the provision of food, and with this in mind we need to be aware of the safety critical dimensions of the context and its relationship to the way that one might develop software systems for such a setting [15]. It is key that people understand the multi-facetted dependent nature of any systems that weaves together online services, autonomous systems, robotics and IoT into the real world, particularly when these are use used in such a way to enhance and support the mutual bond that exists between animals and humans in a domestic setting.

In this setting we are responsible for looking after and caring for animals, passing on parts of this responsibility to an autonomous system means that we are reliant upon the said system doing its job, whether this is opening a cat-flap, feeding or intelligent health-based interventions – we are *entrusting* this role to a technical system. Baier [1] writes, "*the custody of these things that matter to me must often be transferred to others, presumably to others I trust. Without trust, what matters. . .would be unsafe*", which, in the context of this philosophical text means that we transfer trust, we *entrust*, give away control and responsibility to others we reason about it (in our case we are interested this trust *transfer* to intelligent robotic systems), without this sort of "*trust*" and reasoning about potential outcomes there is a risk (things are "*unsafe*") to the things that we care about, things that matter. Negotiating this *in potentia* 'risk' is difficult when one starts to think about the subtle ways that humans interact with companion animals, the subtle personal nature of this interaction (*we know our companion animals*) and the tacit ways that this occurs. How can we design for 'Trust' in this context, when even at a high-level this appears to be so situationally dependent and how do we inform a system about the idiosyncrasies of these interactions? This notion leads us on to think about "*what matters*", the value of "*what matters*" and how we invest in decisions and reason about entrusting human responsibility to technical systems in the context of ACI. In order to address this, we need to start to engage with a range of issues relating to safety, prioritisation, provenance and importance as we start to navigate around the *zoo-socio-technical* landscape when designing trustworthy autonomous systems. In such cases we need to start to think about the reliability of a system - security, how acceptance occurs, the visibility, intelligibility and the explainability of a system, both in respect of ACI and HCI.

Trust and belief are foundational, forming the basis for observable mutual negotiations and interaction between humans [12] and between humans and animals [7]. In the context of this research, we will start to unpack the ways in which humans and animals engage with autonomous technologies, and the role of trust in these settings. The addition of autonomous systems adds another dimension and underlying complexity to these interactions as illustrated by the ongoing research explorations and approaches to design that examine Trust, AI, and Autonomy in a 'mundane' setting [13], but how might we approach this from an ACI perspective and what areas of Trust might we focus upon to inform researchers, policy/lawmakers, and designers? Are there approaches that allow AI to be trained from the perspective of an animal [5], or can we use ethnographic and ethnomethodological approaches [18] to develop a richer more detailed appreciation and understanding of a given context, or could art be used as a design mechanism to engage people with autonomous systems and ACI [19]? Even if we were to understand issues relating to Trust how can we use our findings to inform and generalize design in a multi-species setting?

1.2 Growing a community: Autonomous Systems, Trust & ACI

We would like to encourage people to take part in the workshop who might not normally conder themselves part of the ACI community, but who have insights and understandings, and have developed approaches that can shed light on the role of Trust in designing autonomous systems. In order to do this, we will bring together a range of researchers, artists, and designers to explore how we might deal with intelligent systems in a more animal-centric way, where the social setting and multi-species approaches to designing technology can help us to develop ways to approach RRI and appreciate issues pertaining to agency and start to develop participatory approaches to design and how we might do this in a responsible beneficial way. The notion of participation in design has been previously discussed in the literature [11] [14] and we will expand upon this to further develop responsible approaches to design, particularly in the context of Trust, ACI and autonomous systems. We are at a point where the implementation of a range of autonomous technologies, from robotic systems in the home, to robot toys and feeders (*which we refer to earlier*) and autonomous speech-based systems which can intelligently change the voice of the system (perhaps mimic the sound of a human or animal). It is these technologies that are leading to the development of a range of technology-based enrichment activities and services (as have been previous explored in the ACI research domain *see* [21] http://www.zoojam.org., which could positively impact upon a whole range of ACI contexts, but how do we get people to engage in these debates and thee issues

that surround them, how can we effectively design for such complex contexts in the real world and can the Arts play a role in these debates? This workshop aims to look at some of these interdisciplinary issues and asks people think about how we approach a fast-approaching world, where autonomous-intelligent systems are becoming the norm, where such systems may already be embedded into our lives in ways that we don't see. To encourage different communities to take part in the workshop we are taking an innovative approach and asking for submissions that may take to form of a presentation, case study, pictorial, film, or design. Our focus is to be as open as possible and to encourage participation and encourage creative thinking in order that we might take a wider more holistic perspective in the workshop. We will use social media to promote the event as well as project mailing lists.

1.3 Activities

The workshop will be in-person hosted at the Co-Bot Maker Space, University of Nottingham. This will give attendees the opportunity to directly engage with the technologies that we will be discussing in the workshop, a tour of the lab will take place at midday. This will also allow participants the time to network and explore. The facility includes state-of -the-art digital display technologies and we will create an online space where online participants can engage with the presentations at the workshop.

We will introduce the workshop, this will be followed by a Keynote presentation by Dr Jane Tyson from RSPCA Research, this will start the day and open the workshop creating a space where participants can discuss ideas relating to the workshop theme, a Miro board will allow participants to add comments throughout the sessions.

The day will consist of two separate sessions. Session One (AM) will consist of presented materials (*it is up to the presenter how they would like to do this*), but we envisage that as in previous workshops each person will discuss or show their work and that this will be the catalyst for discussion and debate – after each presentation people will be able to ask the presenter questions (in-person or online). The presentations will be added to the webpage with a small abstract and related materials. The environment will be supportive, and people will be encouraged to offer constructive criticism. The last part of Session One will consist of an introduction to a Responsible Research and Innovation (RRI) tool (a card-set). This will be used as a way to segue into Session Two.

Session Two (PM) will be hosted by the Cat Royale project, an ongoing artistic exploration of cat-human-robot interaction, being led by the artists Blast Theory as part of the UKRI Trustworthy Autonomous Systems Hub. They will share our story so far, engaging with questions of responsibility, trust and autonomy while inviting other participants to reflect on this with a view to establishing an agenda for future research. This session will also engage attendees in an RRI design activity and will allow attendees to understand and discuss the design and deployment of artist-led systems in the wild and how to best engage people in discussions pertaining to the future application of technologies in multi-species settings to understand trust. We will use the Miro board to document this and use the documentation to further discuss the project at the UKRI Trustworthy Autonomous Systems Hub symposium in 2023. We

will end the session by inviting the participants to comment and summarize, offering insights into the workshop.

1.4 Outputs

We envisage several outputs from the workshop, which will include the initial development of a community with a focus on the development and design of trustworthy autonomous systems relating to ACI. As part of the workshop attendees will be able to engage with the artists Blast Theory and discus their work Cat Royale, which will be performed in 2023. As part of the workshop, we will establish a research agenda for the development of autonomous and domestic ACI systems and this will result in the creation of a network which will consist of a community of researchers. We will identify joint research interests and raise possible funding opportunities. As we are developing a range of multi-disciplinary understandings of autonomous ACI systems we will add these to our webpage in order that other researchers can benefit from the workshop. Our webpage is currently hosted at Designing for Trust Autonomous Animal-Centric Robotic & AI Systems.

1.5 Call for Participation

This workshop focuses upon the role that 'Trust' plays in autonomous animal-centric robotic systems and the ways in which multi-disciplinary researchers can further understand how to design, develop, and evaluate such systems by taking a Responsible Research and Innovation (RRI) approach. Understanding, designing for, and negotiating 'Trust' is complex, particularly in contexts where animals, humans, and intelligent systems (including Robotics, AI, & IoT) come together in a social context, in which bonds are created, friendships develop, and mutual care plays a part in the relationship between people and animals. We would welcome submissions from a range of disciplines (as mentioned, but not confined to the themes of the workshop). These may be as presentations, design fictions, case studies, pictorials, ethnographies, artworks, films, and audio-based works. These will be hosted on our webpage as an output of the 'in person' workshop. This will act as a catalyst to start discussion.

ACKNOWLEDGMENTS

REFERENCES

[1] Annette Baier. 1986. Trust and Antitrust. Ethics, 96(2), 231–260. http://www.jstor.org/stable/2381376

[2] Alan Chamberlain, Andy Crabtree, Tom Rodden, Matt Jones, and Yvonne Rogers. 2012. Research in the wild: understanding 'in the wild' approaches to design and development. In Proceedings of the Designing Interactive Systems Conference (DIS '12). Association for Computing Machinery, New York, NY, USA, 795–796. https://doi.org/10.1145/2317956.2318078

[3] Karen Clarke, Gillian Hardstone, Mark Rouncefield, Ian Sommerville (Editors) (2006) Trust in Technology: A Socio-Technical Perspective: 36 (Computer Supported Cooperative Work, 36), Springer.

[4] Andy Crabtree, Alan Chamberlain, Mark Davies, Kevin Glover, Stuart Reeves, Tom Rodden, Peter Tolmie, and Matt Jones. 2013. Doing innovation in the wild. In Proceedings of the Biannual Conference of the Italian Chapter of SIGCHI (CHItaly '13). Association for Computing Machinery, New York, NY, USA, Article 25, 1–9. https://doi.org/10.1145/2499149.2499150

[5] Kiana Ehansi, Hessam Bagherinezhad, Joseph Redmon, Roozbeh Mottaghi, Ali Farhadi (2018). Who Let the Dogs Out? Modelling Dog Behavior from Visual Data. 4051-4060. DOI 10.1109/CVPR.2018.00426.

[6] Fiona French, Ilyena Hirskyj-Douglas, Heli Väätäjä, Patricia Pons, Sabrina Karl, Yoram Chisik, Eleonora Nannoni, Anna Zamansky, Mandeep Mangat, Patrizia Paci, Haruka Kasuga, and Dana Vilker. 2021. Ethics and Power Dynamics in Playful Technology for Animals: Using speculative design to provoke reflection. In Academic Mindtrek 2021 (Mindtrek 2021). Association for Computing Machinery, New York, NY, USA, 91–101. https://doi.org/10.1145/3464327.3464366

[7] David Goode. 2006. Playing with My Dog Katie: An Ethno-Methodological Study of Canine/Human Interaction, Purdue University Press

[8] Haruka Kasuga and Yuichiro Ikeda. 2020. Gap between Owner's Perceptions and Dog's Behaviors Toward the Same Physical Agents: Using a Dog-like Speaker and a Humanoid Robot. In Proceedings of the 8th International Conference on Human-Agent Interaction (HAI '20). Association for Computing Machinery, New York, NY, USA, 96–104. https://doi.org/10.1145/3406499.3415068

[9] Haruka Kasuga, Daisuke Sakamoto, Nagisa Munekata, and Tetsuo Ono. 2017. A Social Robot in a Human-Animal Relationship at Home: A Field Study. In Proceedings of the 5th International Conference on Human Agent Interaction (HAI '17). Association for Computing Machinery, New York, NY, USA, 61–69. https://doi.org/10.1145/3125739.3125759

[10] D. v. d. Linden, B. I. Davidson, O. Hirsch-Matsioulas and A. Zamansky, "On the role of technology in human-dog relationships: a future of nightmares or dreams," in IEEE Transactions on Technology and Society, 2022, DOI: 10.1109/TTS.2022.3207991

[11] Clara Mancini. 2013. Animal-computer interaction (ACI): changing perspective on HCI, participation and sustainability. In CHI '13 Human Factors in Computing Systems (CHI EA '13). Association for Computing Machinery, New York, NY, USA, 2227–2236. https://doi.org/10.1145/2468356.2468744

[12] Christian Meyer. 2019. Ethnomethodology's Culture, in Human Studies; Dordrecht Vol. 42, Iss. 2, 281-303. DOI:10.1007/s10746-019-09515-5

[13] Matthew Pilling, Paul Coulton, Tom Lodge, Andy Crabtree, and Alan Chamberlain. 2022. Experiencing Mundane AI Futures, In the Design Research Society Conference 2022 - DRS2022: Bilbao, 25 June - 3 July, Bilbao, Spain. https://doi.org/10.21606/drs.2022.283

[14] Charlotte Robinson, Ilyena Hirskyj-Douglas, Patricia Pons. 2018. Designing for Animals: Defining Participation, International Conference on Animal-Computer Interaction - Georgia Tech., Atlanta, United States - https://participationaci.wordpress.com, webpage accessed July 2022.

[15] Ian Sommerville. 2011. Software Engineering, 9th edition, Addison Wesley

[16] Lauren Thornton, Bran Knowles, and Gordon Blair. 2022. The Alchemy of Trust: The Creative Act of Designing Trustworthy Socio-Technical Systems. In 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22). Association for Computing Machinery, New York, NY, USA, 1387–1398. https://doi.org/10.1145/3531146.3533196

[17] Van Patter, L. E., & Blattner, C. 2020. Advancing Ethical Principles for Non-Invasive, Respectful Research with Nonhuman Animal Participants, Society & Animals, 28(2), 171-190. https://doi.org/10.1163/15685306-00001810

[18] David Randall, Mark Rouncefield, and Peter Tolmie (2021). Ethnography, CSCW and Ethnomethodology, Computer Supported Cooperative Work (CSCW), vol. 30, no. 2, pp. 189-214.

[19] Maja Smrekar & Jonas Jørgensen (2022) !brute_force – Soft Resilience, Performed at Ars Electronica, https://ars.electronica.art/planetb/en/brute-force/ (accessed Oct 2022)

[20] Erika Amethyst Szymanski, Robert D. J. Smith & Jane Calvert (2021) Responsible research and innovation meets multispecies studies: why RRI needs to be a more-than-human exercise, Journal of Responsible Innovation, 8:2, 261-266, DOI: 10.1080/23299460.2021.1906040

[21] Zoo Jam - http://www.zoojam.org, webpage accessed July 2022