



## **Agent archetypes for human-drone interaction: Social robots or objects with intent?**

Downloaded from: <https://research.chalmers.se>, 2024-04-25 11:35 UTC

Citation for the original published paper (version of record):

Baytas, M., La Delfa, J., Ljungblad, S. et al (2020). Agent archetypes for human-drone interaction: Social robots or objects with intent?. CEUR Workshop Proceedings, 2617

N.B. When citing this work, cite the original published paper.

---

# Agent Archetypes for Human-Drone Interaction: Social Robots or Objects with Intent?

**Mehmet Aydın Baytaş**  
Chalmers University of  
Technology  
Gothenburg, Sweden  
baytas@chalmers.se

**Joseph La Delfa**  
RMIT University  
Melbourne, Australia  
joseph@exertiongameslab.org

**Sara Ljungblad**  
University of Gothenburg  
Gothenburg, Sweden  
sara.ljungblad@gu.se

**Morten Fjeld**  
Chalmers University of  
Technology  
Gothenburg, Sweden  
fjeld@chalmers.se

## Abstract

Departing from our earlier work on conceptualizing “social drones,” we enrich the discussion using notions of “agent archetypes” and “objects with intent” from recent interaction design literature. We briefly unpack these notions, and argue that they are useful in characterizing both design intentions and human perceptions. Thus they have the potential to inform the creation and study of HDI artifacts. Upon these notions, we synthesize relevant implications and directions for design research, in the form of research questions and design challenges. These questions and challenges inform our current and future work. We submit our resources, arguments, aims, and hypotheses to the iHDI 2020 community as a reflective exercise, aiming to refine our work in synergy with other participants.

## Author Keywords

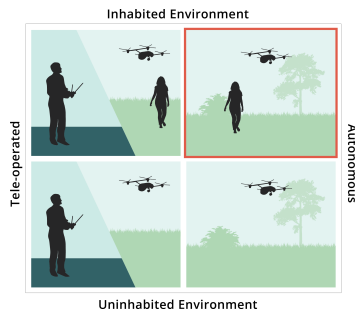
Autonomous drones; design philosophy; design theory; drones; human-drone interaction; social drones; unmanned aerial vehicles.

## Introduction

In previous work, we had proposed the term *social drones* to describe applications where autonomous drones operate in human-populated environments [2, 3]. Here, our choice of the term *social* was inspired by a particular definition for “social animals” as beings that “regulate each

---

This paper is published under the Creative Commons Attribution 4.0 International (CC-BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.  
*Interdisciplinary Workshop on Human-Drone Interaction (iHDI 2020)*  
*CHI '20 Extended Abstracts, 26 April 2020, Honolulu, HI, US*  
© Creative Commons CC-BY 4.0 License.



**Figure 1:** We had proposed the term *social drones* to cover autonomous drones operating in human-populated environments. (Figure from [2].)

other’s nervous systems” [6]. Departing from this definition for sociality, we made the observation that some form of social/regulatory interaction between any two living agents is unavoidable when they occupy the same space and can observe each other. For example, a cat and a human in the same room, given enough time, will regulate the affect and behavior of each other in various ways. Similarly, we argued that an autonomous embodied agent in an inhabited space can be described as *social*. Thus we meant to imply two things:

1. It is unavoidable that a flying machine in the same space will affect any humans present. Thus, human factors must be foregrounded in the design of autonomous drones operating in human-populated environments.
2. A social drone must have capabilities and present affordances to capture human input. If the drone does not perceive and respond to the human (i.e. be regulated by the human), the design risks being perceived as “antisocial” and undesirable.

This identification of autonomous drones in human-populated environments as a distinct category of HDI has been fruitful in scaffolding our work.<sup>1</sup> However, the term “social” has also turned out to be problematic, in that it evokes mental models and expectations grounded in consciousness and sentience. This can be an issue for users; and also for designers, as it can limit the design space, perhaps unnecessarily.

In this position paper, we depart from our previous conceptualization of “social drones” and consider Rozendaal,

<sup>1</sup> See: [wasp-hs.org/projects/the-rise-of-social-drones-a-constructive-design-research-agenda/](http://wasp-hs.org/projects/the-rise-of-social-drones-a-constructive-design-research-agenda/)

Boon, and Kaptelinin’s (2019) analysis of agent archetypes in HCI [15]. In doing so, we wish to move beyond *social drones* as a category that is limited to archetypal *social robots*. Rather, we highlight that much of the work that is relevant to this space (including ours) spans, in addition to *social robots*, categories like *ambient agents* and *objects with intent* (OWI). Further, we make the case that there are numerous situations where it will serve designers to explicitly prefer non-anthropomorphically grounded archetypes to scaffold mental models. These situations may include, but are not limited to, safety-critical and professional applications such as search and rescue, fire response, construction, etc. – where the correctness of both users’ and bystanders’ mental models might be consequential.

In what follows, we first introduce the “agent archetypes” analysis and the four relevant categories of agents that might figure in scaffolding HDI design work, drawing heavily on Rozendaal et al. [15]. We then discuss the implications of applying such an analysis in HDI research, with a focus on open questions and related hypotheses. Our plan is to explore these implications in our own future work, through creating and studying HDI via constructive design research. We are publishing these discussions as a reflective exercise, in order to explore opportunities for synergy with other participants at the Interdisciplinary Workshop on Human-Drone Interaction (iHDI 2020) [4].

## Agent Archetypes

In their 2019 article, in order to unpack the OWI concept used to scaffold their design work, Rozendaal et al. present an analysis where they cluster four “agent archetypes” relevant for computing artifacts [15]. The first category here is that of *ambient agents*, which appear as part of “ambient intelligent environments” [1]. In principal, ambient agents are components that sense, interpret, and actuate

	Grounding Metaphor	Interaction: Explicit vs. Implicit	Interaction: Direct vs. Semantic
<b>Ambient Agents</b>	Environment	Implicit	Direct
<b>Conversational Agents</b>	Human	Explicit	Semantic
<b>Social Robots</b>	Being	Flexible	Semantic
<b>Objects with Intent</b>	Thing	Flexible	Direct
<b>Non-agents</b>	Thing	Explicit	Direct

**Table 1:** Comparing agent archetypes. Based on [15], with the addition of the category of *non-agents*; to show – within the same analysis framework – how artifacts meant *not* to evoke a sense of agency differ from “agent” artifacts.

the environment, thus being “experienced collectively as a supportive ambient intelligent presence in the environment” [15]. The second category is *conversational agents* that “rely on natural language to interact with humans through written text or speech” [15]. These may be implemented as parts of GUIs, virtual characters, within physical artifacts, or through instant messaging interfaces. Third, the analysis exposes the category of *social robots*, which are physical “mechatronic agents.” Often, these are designed with humanoid or animal-inspired forms. The authors note that studies with such robots indicate that their “intelligence” may often be “overestimated,” since people’s expectations may be influenced by their experiences with living beings.

The central topic in Rozendaal et al.’s work is the category of OWI, which describes artifact designs that exploit “the meaning of everyday things as the site for their intelligence and agency” [15]. Thus, the OWI concept can scaffold interaction designs meaning to evoke a sense of “collaborative partnership” between the user and the thing, while avoiding issues such as overestimation, uncanniness [12], and over-attachment [13].

Table 1 summarizes the key characteristics of the four agent archetypes explained above. We also add a fifth category of *non-agents*, which shows – within the same analysis frame-

work – how artifacts meant *not* to evoke a sense of agency differ from “agent” artifacts.

## Implications and Directions

We would argue that the agent archetypes framework has significant implications in terms of how it might inform the creation and study of HDI designs. Here, we propose a number of topics and directions for HDI research where such analysis may be fruitful. We intend to adopt some of these proposals as research questions and design challenges to direct our own future work. We hope that we will also find other participants at iHDI 2020 who have interest in these topics, and some of our future efforts may ensue in synergy.

*What are the places for different agent archetypes in HDI?* Rozendaal et al.’s analysis is useful for characterizing different agent archetypes within the broader contexts of product design, human-agent interaction (HAI), and human-robot interaction (HRI). However, this framework is not prescriptive in the sense that it might tell us where and when each archetype might be useful – particularly in the context of HDI. Relevant open questions include:

- Where and when is it desirable to design to embody particular agent archetypes? What are some specific use cases where each agent archetype might be more appropriate than the others?
- What happens if the agent type is inappropriate for the context or use case? What might be some “modes of failure” that relate to agent archetypes, and how might we trace them back to their cause?

In response to these questions, for example, we hypothesize: in HDI, Objects with Intent and Ambient Agents archetypes (and “non-agents”) may be more relevant and/or desirable over Social Robots in safety-critical and professional contexts (e.g. search and rescue, fire response, construction) where the correctness of both users’ and bystanders’ mental models might be consequential.

#### *How might we embody agent archetypes in HDI?*

Though the agent archetypes framework itself is not prescriptive, there exists ample literature with theory, tools, and exemplars that can scaffold design work based on any one of the archetypes.<sup>2</sup> Focusing on HDI, we note an abundance of such resources that could support designing drones as Social Robots,<sup>3</sup> but we are not aware of any resources which might inform HDI designs based on the Objects with Intent concept. Thus, for HDI, we might pose the following open questions:

- To what extent is the designer even in control of how the interaction artifact will be perceived? How strongly

<sup>2</sup>Ambient Agents, Conversational Agents, and Social Robots are now canonical topics in the relevant literatures. For Objects with Intent, see: [14, 15, 16, 17]

<sup>3</sup>For reviews, see: [2, 10]

are designers’ intentions and human perceptions correlated, with respect to the agent archetypes analysis, in the context of HDI? How stable are these perceptions, between different populations of users and bystanders, and across time?<sup>4</sup>

- How might we create frameworks, tools, and strategies based on agent archetypes to expedite HDI designs?
- Would it be sensible to design HDI agents that may ‘switch’ the archetype they embody, depending on the context?

## **Critical Discussion**

As an additional point for discussion, we believe that the particular design choices and human perceptions related to agent archetypes may in fact remain inconsequential, *as long as the artifact is working fine*. We thus hypothesize that the relevance of agent archetypes is amplified when the agent is not behaving as we expect it to believe. Invoking terminology from the literature;<sup>5</sup> our argument is that design choices and perceptions around agent archetypes are less consequential when the agent-artifacts are *ready-to-hand*, and they become consequential when the agent-artifact becomes *present-at-hand*.

Furthermore: agent archetypes and OWI are relatively new ideas, and here we have based our thoughts on one particular reading of them. Other interpretations may be possible. Our reading focuses on the comparative classification expressed on Table 1. Specifically, departing from the summarization of OWI as designs that exploit “the meaning of

<sup>4</sup>See: [11]

<sup>5</sup>Our terminology is based on Dourish’s unpacking [5] of Heidegger’s phenomenology [7]; an unpacking that draws on earlier work by Winograd and Flores [18].

everyday things as the site for their intelligence and agency" [15]: the way we understand the phrase "everyday things" has been not as a synonym for "familiar objects," but as just "objects" as opposed to *animate* beings. This relates to how the "grounding metaphors" (Table 1) for different agent archetypes compare. However, while we focus on "every-day *things*," we acknowledge that another reading may find value focusing on "*everyday* things." Thus, the idea of toys, furniture, and other familiar everyday objects could turning agents – and even being perceived to have intelligence – can become a design resource.

## Conclusion

In this position paper, we aimed to capture the notions of "agent archetypes" and *Objects with Intent* which came to our attention through work published by Rozendaal et al. [15], and bring these to the attention of the iHDI 2020 community. Our discussion departs from our earlier work in conceptualizing *social drones* [2, 3], and aims to move beyond this conceptualization towards design resources that might serve a broader variety of contexts and use cases. Building on ideas in the literature, we synthesized a number of implications and directions for design research, in the form of research questions and design challenges. We have already been engaged in design research efforts that, albeit indirectly, relate to these ideas [9, 8]. In future work, we hope to address some of these questions and challenges more directly. We welcome critiques and contributions from the iHDI 2020 community towards this agenda.

## REFERENCES

- [1] Emile Aarts and Reiner Wichert. 2009. Ambient intelligence. In *Technology guide*. Springer, 244–249.
- [2] Mehmet Aydın Baytaş, Damla Çay, Yuchong Zhang, Mohammad Obaid, Asım Evren Yantaç, and Morten Fjeld. 2019. The Design of Social Drones: A Review of Studies on Autonomous Flyers in Inhabited Environments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 250, 13 pages. DOI: <http://dx.doi.org/10.1145/3290605.3300480>
- [3] Mehmet Aydın Baytaş, Mohammad Obaid, Joseph La Delfa, Asım Evren Yantaç, and Morten Fjeld. 2019. Integrated Apparatus for Empirical Studies with Embodied Autonomous Social Drones. In *1st International Workshop on Human-Drone Interaction*. Ecole Nationale de l'Aviation Civile [ENAC], Glasgow, United Kingdom. <https://hal.archives-ouvertes.fr/hal-02128387>
- [4] Mehmet Aydın Baytaş, Markus Funk, Sara Ljungblad, Joseph La Delfa, and Florian 'Floyd' Mueller. 2020. iHDI 2020: Interdisciplinary Workshop on Human-Drone Interaction. In *Proceedings of the 2020 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '20)*. ACM, New York, NY, USA.
- [5] Paul Dourish. 2004. *Where the action is: the foundations of embodied interaction*. MIT press, Chapter 4.
- [6] Lex Fridman. 2018. Lisa Feldman Barrett: How the Brain Creates Emotions | MIT Artificial General Intelligence (AGI). <https://youtu.be/qwsft6tmvBA>. (2018). Accessed: 2019-02-11.
- [7] Martin Heidegger, John Macquarrie, and Edward Robinson. 1962. Being and time. (1962).

- [8] Joseph La Delfa, Mehmet Aydın Baytaş, Rakesh Patibanda, Hazel Ngari, Rohit Ashok Khot, and Florian “Floyd” Mueller. 2020. Drone Chi: Somaesthetic Human-Drone Interaction. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. ACM, New York, NY, USA.
- [9] Joseph La Delfa, Mehmet Aydın Baytaş, Olivia Wichtowski, Rohit Ashok Khot, and Florian Mueller. 2019. Are Drones Meditative?. In *Proceedings of the 2019 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '19)*. ACM, New York, NY, USA. DOI : <http://dx.doi.org/10.1145/3290607.3313274>
- [10] Chun Fui Liew and Takehisa Yairi. 2020. Companion Unmanned Aerial Vehicles: A Survey. (2020).
- [11] Sara Ljungblad, Jirina Kotrbova, Mattias Jacobsson, Henriette Cramer, and Karol Niechwiadowicz. 2012. Hospital Robot at Work: Something Alien or an Intelligent Colleague?. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. Association for Computing Machinery, New York, NY, USA, 177–186. DOI : <http://dx.doi.org/10.1145/2145204.2145233>
- [12] Masahiro Mori, Karl F MacDorman, and Norri Kageki. 2012. The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine* 19, 2 (2012), 98–100.
- [13] Byron Reeves and Clifford Ivar Nass. 1996. *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge university press.
- [14] Marco Rozendaal. 2016. Objects with Intent: A New Paradigm for Interaction Design. *Interactions* 23, 3 (April 2016), 62–65. DOI : <http://dx.doi.org/10.1145/2911330>
- [15] Marco C. Rozendaal, Boudewijn Boon, and Victor Kaptelinin. 2019. Objects with Intent: Designing Everyday Things as Collaborative Partners. *ACM Trans. Comput.-Hum. Interact.* 26, 4, Article Article 26 (June 2019), 33 pages. DOI : <http://dx.doi.org/10.1145/3325277>
- [16] E. van Beek. 2017. *What does it have in mind?: Collaborating with guide dogs, backpacks and Objects with Intent*. Master's thesis. TU Delft.
- [17] F.A.D. Van Boheemen. 2016. *Diem: An animated house that cares with Intent*. Master's thesis. TU Delft.
- [18] Terry Winograd, Fernando Flores, and Fernando F Flores. 1986. *Understanding computers and cognition: A new foundation for design*. Intellect Books.