



Delft University of Technology

Grasping AI experiential exercises for designers

Murray-Rust, Dave; Lupetti, Maria Luce; Nicenboim, Iohanna; Hoog, Wouter van der

DOI

[10.1007/s00146-023-01794-y](https://doi.org/10.1007/s00146-023-01794-y)

Publication date

2023

Document Version

Final published version

Published in

AI and Society

Citation (APA)

Murray-Rust, D., Lupetti, M. L., Nicenboim, I., & Hoog, W. V. D. (2023). Grasping AI: experiential exercises for designers. *AI and Society*. <https://doi.org/10.1007/s00146-023-01794-y>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.



Grasping AI: experiential exercises for designers

Dave Murray-Rust¹ · Maria Luce Lupetti¹ · Iohanna Nicenboim¹ · Wouter van der Hoog¹

Received: 11 May 2023 / Accepted: 18 September 2023
© The Author(s) 2023

Abstract

Artificial intelligence (AI) and machine learning (ML) are increasingly integrated into the functioning of physical and digital products, creating unprecedented opportunities for interaction and functionality. However, there is a challenge for designers to ideate within this creative landscape, balancing the possibilities of technology with human interactional concerns. We investigate techniques for exploring and reflecting on the interactional affordances, the unique relational possibilities, and the wider social implications of AI systems. We introduced into an interaction design course ($n = 100$) nine ‘AI exercises’ that draw on more than human design, responsible AI, and speculative enactment to create experiential engagements around AI interaction design. We find that exercises around metaphors and enactments make questions of training and learning, privacy and consent, autonomy and agency more tangible, and thereby help students be more reflective and responsible on how to design with AI and its complex properties in both their design process and outcomes.

Keywords Artificial intelligence · Design · Prototyping · Design education · Experiential methods · AI exercises · More-than-human design

1 Introduction

Designers increasingly need to develop a facility with artificial intelligence, as it becomes part of the way that products and services function and appears in an increasing number of the contexts in which designers work (Benjamin et al. 2021; Dove et al. 2017). However, there are several challenges for design students in engaging with AI, from the broadness of the term AI and the fuzziness with which it is applied (Littman 2021), to the difficulty of getting to grips with the technical and computational complexities of these systems (Yang et al. 2020; Nicenboim et al. 2022). These challenges around understanding and making sense of the new capabilities of AI become urgent as the technology emerges from

its latest winter into a new spring, developing at a fast pace (Littman 2021; Samoli et al. 2020; Floridi 2020).

The range of techniques for making creative use of AI has been rapidly growing: runway offered easy access to generative spaces and now video (Runway 2020); EdgeImpulse offers sound and gesture classification for microcontrollers with training through a web interface (EdgeImpulse 2019); the current sets of generative image models such as DALL-E, Midjourney and StableDiffusion and language models (ChatGPT, etc.) allow a natural language interaction through the use of prompts. Along with learning materials for more traditional toolkits (TensorFlow 2015; OpenCV 1999) and model development and exchange initiatives (e.g. HuggingFace 2016), these form a downward pressure on the technical barrier to entry, even as the complexity of the underlying models increases. The conceptual barrier can remain high, though, reducing the possibility for designerly engagement and appropriation. There is a large jump from “my first ML model” to understanding the implications of ML technology, and designers often want—and need—to engage with these implications. Creating models in practice helps, but this needs conceptual framing to help direct and contextualise the activities—for example, in related fields, courses such as Creative Applications of Deep Learning (Mital 2016) and the more provocative follow up Cultural Appropriation

✉ Dave Murray-Rust
d.s.murray-rust@tudelft.nl

Maria Luce Lupetti
m.l.lupetti@tudelft.nl

Iohanna Nicenboim
I.Nicenboim@tudelft.nl

Wouter van der Hoog
wvanderhoog@gmail.com

¹ Human Centred Design, Industrial Design Engineering, TU Delft, Delft, The Netherlands

with Deep Learning (Mital 2021) look at visual practice, or Machine Learning for Musicians and Artists (Fiebrink 2022) unpack these systems for creative practitioners.

However, simply thinking about why it is hard for designers to appropriate AI technologies into their practices also misses a key question: what can design practice bring to the development and understanding of AI systems (Benjamin et al. 2021), especially as the technologies become more pervasive and more collaborative (Wang et al. 2020). Designers have their own strategies for making use of, critiquing and appropriating new technologies (Westerlund and Wetter-Edman 2017), so there is an interest in understanding what designerly methods could reveal about human–AI relations, particularly where it involves interactions between humans and technological systems—considering the “social, political, ethical, cultural, and environmental factors of implementing AI into daily human-to-computer interactions” (Wong 2018). Design research methods, speculations (Auger 2013; Kirman et al. 2022), fictioning (Forlano and Mathew 2014; Wong et al. 2017; Troiano et al. 2021; Benjamin et al. 2023), probes and toolkits (Sanders and Stappers 2014), more than human design (Coulton and Lindley 2019) and the general practices of Research through Design (RtD) (Giaccardi 2019; Stappers and Giaccardi 2017), are all well suited to thinking into the socio-technical aspects (Holton and Boyd 2021; Sartori and Theodorou 2022; Theodorou and Dignum 2020), possibilities, and implication of AI in everyday life, just as they have been applied to understanding digital sensing technologies (Pierce 2021), blockchains (Murray-Rust et al. 2022a), the future of automation (Cavalcante Siebert et al. 2022) and so on.

Our aim is to help students to design products and services that make use of AI technologies, while developing a critical understanding of its implications. This means articulating both the technical and relational aspects of AI so that meaningfully shape the development of products, services and systems even if they are not intimately familiar with the technical details of its operation. As such we are looking for ways to sensitise interaction designers to AI, to create experiences rather than explanations. In relation to the typology developed by Yang et al. (2020) of ways to aid designers around AI, our work contributes to the early stages of ‘creating AI-specific design processes’ by probing with concrete exercises ways in which educators can support designers in ideating in a space mediated by the capabilities and implications of AI systems.

To explore this space, we created a set of methods for designing AI driven products and services (Sect. 3.2) that draw on theories about how people relate to technology and AI in particular. These methods take the form of short, autonomous, experiential exercises that can be used to develop and enrich the design of interactive technological products and services. We introduced these exercises

partway through an interaction design course (Sect. 3.1), where students ($n = 100$) in small groups ($n = 28$) are asked to design future products and services through iterative prototyping and testing. We collected an immediate written reaction from each group as to what the students had done with the materials, and the aspects they found useful or resonant. We interviewed a self-selecting subset of the students ($n = 12$) and their coaches ($n = 7$) to dig deeper into questions of how the methods had changed their understandings and relations to AI.

To explore the potential of these methods, we investigate the following research questions:

- *RQ1*: How do the exercises stimulate and modulate changes to the students’ design process to accommodate AI, in particular the way that they are conceptualising and prototyping their projects?
- *RQ2*: How do these experiential exercises affect student’s grasp of AI and ML, in particular in relation to interactional, relational and contextual qualities which are key points in the recent theoretical developments in AI within HCI?
- *RQ3*: How do the exercises help to develop a critical design perspective while engaging with AI technology as a socio-technical system?

Through investigating and discussing these research questions, the contributions of this work are:

1. A set of exercises that translate theoretical developments in design and AI, into experiential exercises for designers that can be carried out autonomously, with reflection on the experiential, pragmatic and reflective qualities that made the exercises effective. These exercises are available at [redacted] for future use and development.
2. Insights into how and for what to apply these exercises in a pedagogical context to support design processes for creating AI enabled products and services.
3. Insights about how these exercises affected student’s reasoning and design activities, bringing agency, relationality and criticality alongside development of technical facility.
4. Methodological reflections around the possibilities afforded by the methods and how these contribute to nurturing a uniquely designerly AI culture that supports future design education.

2 Background

Working with AI presents particular challenges for designers. One of them is around engaging with emerging and complex technologies, with different behaviours from traditional

design materials. Yang et al. (2020) point at two key challenges: the uncertainty about the capability of AI systems and the complexity of their outputs. The second challenge is around understanding AI, given that the metaphors and imaginaries around it, obscure the real processes that are needed for maintaining such a technology (Murray-Rust et al. 2022b). Contrary to competing terms like ‘complex information processing systems’ or ‘machine intelligence,’ the term AI fires the mind with ideas of human-like reasoning. While these imaginaries seem to be better for marketing, they are certainly no good for developing a grounded sense of the capabilities of AI as a technology Hildebrandt (2020).

2.1 Designing AI

Despite the challenges for designers to engage with AI, there are currently many areas where design and AI touch on each other.

At a low level, there is growing attention to the meeting of AI and user experience (UX), as the new possibilities offered by the technology allow new kinds of interaction, and are susceptible to new pitfalls. Techniques are emerging that help to create user interfaces that work with AI systems (Subramonyam et al. 2021a, b), or support innovating AI-powered services and systems within enterprises (Yildirim et al. 2022). This can be seen in Microsoft’s guidelines for human–AI interaction (Amershi et al. 2019), or Google PAIR’s Guidebook (PAIR 2020), as well as efforts to bring HCI together with AI (Inkpen et al. 2019). Recently, the identification of ‘AI capabilities’ (Yildirim et al. 2023) provides a concrete way to think about design spaces for the interactional aspects of computational system. Negotiation between AI and HCI can be deep and subtle: interactional affordances help to calibrate trust and reliance between humans and AI; conceptual metaphors sculpt the relations formed with conversational agents (Jung et al. 2022; Khadpe et al. 2020); and appropriate abstractions make AI qualities at hand for creative practitioners (Fiebrink 2019; Tremblay et al. 2021). User experience, in its broader sense, goes beyond designing the immediate experiences, with work starting to consider how to develop frameworks for creating more or less personal, dependent and discretionary interfaces (Kliman-Silver et al. 2020), or at how to generate heuristic models of meaningful engagement with AI artworks (Hemment et al. 2022b).

Zooming out slightly, a collection of theoretical issues around AI relate to emerging fields in the third (and fourth) wave HCI and the philosophy of design communities. Scholars in those areas have grappled with how the concepts used in design and HCI practices might be tied to the industrial era, and how they might have to change and adapt to the new kinds of products and materials which are enabled by AI. This includes lines of research such as post-industrial

design, more than human practices (Giaccardi and Redström 2020), entanglement thinking (Frauenberger 2020; Murray-Rust et al. 2019; Hodder 2016), fluid assemblages and multi-intentionality (Redström and Wiltse 2018; Wiltse 2020). All these offer vibrant pictures of a new set of relationships between humans and the material world in which both entities ‘co-constitute’ each other. Along with reorienting the relationships between humans and non-humans, scholars within those fields have been rethinking what it is to ‘do design’, breaking with traditions focused on the subject–object dichotomy (Giaccardi and Redström 2020), where design goes beyond a mere problem solving enterprise and becomes an ongoing and more inclusive practice. Although these theoretical developments seem to be gathering momentum, they are still not fully translated into practical tools for designers—the jump from Barad’s agential realism (Barad 2007) to configurations of bits and programmes takes careful work (Scurto et al. 2021; Seymour et al. 2022; Sanches et al. 2022).

At a broader scale, beyond the immediate interactions, some of the theories and practices at play are oriented towards engineering particular system qualities and properties: value-sensitive design can help to make sense of fluid and evolving systems (de Reuver et al. 2020), where many different human values may be at play (Yurrita et al. 2022; Fish and Stark 2021; Shen et al. 2021); questions of meaningful human control modulate the relations of responsibility between humans and automated systems (Cavalcante Siebert et al. 2022), as does responsible AI design (Benjamins et al. 2019). Here, design is an instrumental part of making systems behave in certain ways. AI ethics is a broad field (Hagendorff 2020), and well as directly affecting system properties, work from the Fairness Accountability and Transparency (FAccT) community looks to support documentation that helps maintain these properties in communities, such as documentation for models and datasets (Mitchell et al. 2019; Gebru et al. 2018), and the ethical aspects of system development (Mohammad 2021; Murray-Rust and Tsiakas 2022).

2.2 AI and design education

The specificities and challenges of AI and ML technologies add up to an ongoing discourse at the intersection of design education and technological progress. Traditional formats and scopes for carrying out design are being questioned and revised, with canonical, linear, causal and instrumental approaches being criticised in favour of novel models inspired by complexity theory, system science, and practical philosophy. This moves towards an aim of reconceptualising design as a moral act (Findeli 2001; Lin 2014). Designers and design researchers, in fact, are increasingly recognised as actors whose decisions have ethical as well

as political implications (Lloyd 2019). Parallel to this, the societal implications of AI and ML are more clearly pervasive and unpredictable. However, when introduced in design education, these technologies are typically either approached as the ultimate tools to learn or used as ‘context’ for grounding alternative and critical design explorations. On the one hand, ML courses are increasingly offered to design students for promoting ML/AI literacy but remain an addition to the main curricula, rather than being integrated into project courses [as in (Jonsson and Tholander 2022; van der Vlist et al. 2008)]. These technologies are still rarely integrated in design education (Dove et al. 2017), and often approached with the belief that even just exposing students to cutting edge technology can stimulate the emergence of innovative and technologically advanced design solutions to real problems (McCardle 2002). In other cases, however, the approach is diametrically different: AI and ML are conceptualised as phenomena to be understood and questioned, because of its potential impact in society. For instance, Auger (2014) used the theoretical lens of domestication for challenging students to ideate future domestic robots and reflect on their implications in everyday settings. The two perspectives on AI and ML, the focus of which we summarise as technical competency vs critique, tend to remain distinct approaches with apparently opposite scopes. Even when there is an explicit commitment to bridge the two approaches, existing pedagogy struggles to combine the ambition to build AI literacy while also fostering a critical mindset around AI/ML projects, and reflections do not lead to rich critiques about situated and contextual implications of AI and ML unless they are integrated into project development. For some counterexamples, Jonsson and Tholander (2022) purposefully crafted a course for students to approach and appreciate AI tools as creative partners, and learned that AI qualities, such as uncertainty, imperfection, and underdetermination, can be a rich source of inspiration for generating creative expressions as well as powerful triggers of reflection. Mital’s Cultural Appropriation with Deep Learning course (Mital 2021) weaves together learning about the operation of deep networks with recognising their role in society. Fiebrink’s work (Fiebrink 2019) distinctively looks at ML as a design material and situates it within project development. Perhaps most similar to the work outlined here is ‘Graspable AI’ (Ghajargar et al. 2021, 2022; Ghajargar and Bardzell 2023), which brings together tangibility and AI, using explanation as a path to understanding and form as a language for communicating AI affordances. Even in these cases, however, the emphasis is on one side of the spectrum, that is on how to teach ML effectively to any population and enable the emergence of new creative outputs.

The disciplinary call for exposing the design questions involved in making AI and ML systems—as well as the complexity and trade-offs that implementing these in the world

implies (Bilstrup et al. 2022)—remains largely unanswered. Our work sits at the intersections of these experiences and aims to fill the gap between technical efforts and critical explorations. Specifically, we set out to integrate AI and ML explorations within the development of design projects, in a way that both enable students to build AI literacy, as well as to empower them to take a critical stand towards these technologies in society.

2.3 Summary and research direction

Part of the work of design as a discipline is to mediate between these philosophies and actionable practices that can be brought to bear on particular situations. That is the starting point for the work presented in this paper: we are interested in how to bring conceptual developments from design theory and AI into something that is at hand to design students, that can make a difference to how they go about conceptualising and prototyping interactions.

To bridge the gap between the practical and technical engagement with AI, we propose that three levels of engagement between AI and design are all potentially at play within design projects creating AI-powered systems:

Interactional affordances of AI that allow new means of interaction between systems and people. At a low level, AI brings new possibilities for sensing, responding, recognising and classifying from which to build interactions. These interactional affordances and possibilities for action (Stoffregen 2003) offered by machine intelligence can take the form of capabilities offered by the technology (see (Yildirim et al. 2023) for a comprehensive overview), but also of modulations of existing capabilities with AI-specific qualities such as probabilistic outcomes.

AI relationality as it is brought into constellations and forms new relations between people and things. Beyond the immediate interaction, design with AI intervenes conceptually and materially in constellations (Coulton and Lindley 2019) of humans and objects. Designers must navigate the increased agency and depth of interaction that intelligent systems bring, and the changes in the way that we understand and relate to technological systems.

Wider implications of AI as it affects social structures and people’s lives outside the immediate interactions. Concerns about the implications of systems are not new, but AI and data-driven systems that are built through processing large amounts of data about people bring new and subtle ways in which they can be unfair or unjust, more blurring of responsibility and more potential unintended consequences at scale.

For the work at hand, we are interested in how these levels relate to design education, in particular how students start to engage with AI as a design material. To create a broad coverage, we looked at creating methods that could create

engagement with the specifics of working with AI systems on these three levels, as well as balancing the educational concerns of developing a better understanding of and facility with the technology, and encouraging critique. Based on this, we created a set of ‘design exercises’—rapid, experiential engagements that draw on the theoretical developments above but can be carried out productively in the context of conceptual development and prototyping interactions with AI systems (Fig. 1).

3 Study

3.1 Course and context

The context of this study is a one-semester (20 week) design and prototyping course for first year Masters students in the ‘Design for Interaction’ programme at TU Delft. All students in the course ($n = 100$) have a design background with

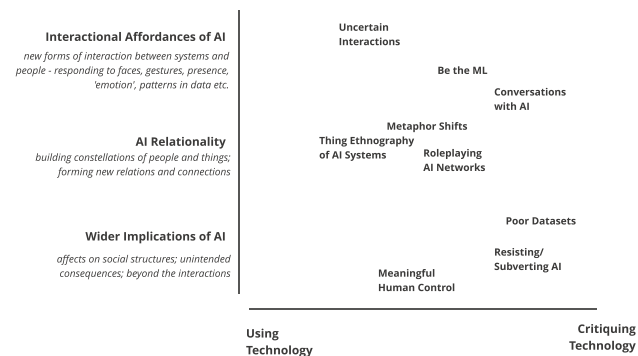


Fig. 1 Situation of the methods underdiscussion across two axes: (i) the level of consideration, from direct interactional affordances, through building relationality out to wider implications and (ii) the balance between developing facility with the technology and critiquing it's uses

a mixed range of computational skills, from no technical knowledge to beginner level in software engineering. The students were grouped by the course coordinators into 28 teams, and coached by seven experienced coaches from the Industrial Design Engineering faculty.

The course is structured in three stages, with student teams of 3 or 4 students working 13 per week on their design project (Fig. 2). They worked on design briefs that asked them to speculate about near future interactions supported by technology. Many of these briefs were provided by client companies, for example new forms of human–vehicle interactions, possibilities for more sustainable cooking through smart kitchen appliances and pervasive computing in hotel rooms. The students had little to no pre-course familiarity with Machine Learning and AI methods, theories and tools; however, most of the coaches had at least some experience with these technologies.

The main learning objective of the course is to introduce students to various ways of prototyping with interactive technology. Students were asked to design within the context of the client company, creating and testing a new iteration of their prototype each week in discussion with their coach. They were prompted to draw on some form of AI or ML, although technical capabilities could be acted out rather than implemented in code. The course ran in three phases: a “First Shot” familiarisation with AI and technology, “Iterating Forward” to develop concepts and “Polishing Up” the final ideas and prototypes (Fig. 2), with an exhibition at the end of each stage. Client companies were invited at the end of each stage to provide feedback to the student teams, organised in the form of an exhibition with interactive prototypes.

3.2 Exercises

The intervention involved a set of 9 exercises (Table 1). Each design exercise was introduced on a single page, containing a title, a short description and instructions on how to execute

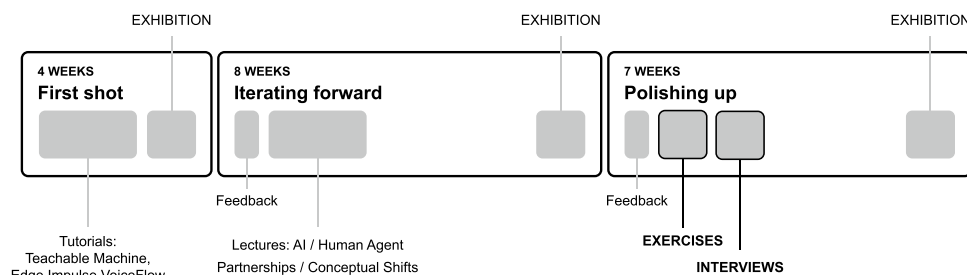


Fig. 2 Course structure; In the first stage (4 weeks) students were given context on AI and ML, and hands-on engagements with AI technology were provided through a series of workshops with existing tools (Edge Impulse, Teachable Machine and Voiceflow). At the end of the first stage, teams presented multiple ideas demonstrated in multiple early prototypes. The second stage introduced lectures cov-

ering AI capabilities, Human–Agent partnerships and the conceptual shifts mentioned above as the students developed their core concept, leading to a second exhibition of interactive prototypes. The third stage introduced the exercises discussed in this paper, as the students refined their projects towards a highly immersive final exhibition with one or more interactive prototypes

the exercise, a background section describing the intent, usefulness and ideas behind the exercise and references to papers and related projects (Fig. 3, see supplementary material for full set). The choice of this set of exercises was exploratory: we derived them from a combination of existing

design practices, emerging work from the researchers and the theories mentioned above, through extensive discussion between the researchers. We aimed to have a spread of exercises across immediate interactional affordances of AI, mid-level human–machine relations and concerns about the

Table 1 Name, key references and description of each method

Method	Description and inspirations
Uncertain interactions	Look through the state diagram of your interaction; for each change of state, imagine replacing it with probabilistic, uncertain or in-between outcomes Bowler et al. (2022); Benjamin et al. (2021)
Be the ML	From a live view of the data inputs to your system, try to perform the activity yourself, then explain what you're doing to someone else Devendorf and Ryokai (2015); Scurto et al. (2021)
Poor datasets	Iteratively remove examples from your dataset, decreasing diversity of the input and retraining the model until something problematic happens. Elwes (2019); Buolamwini and Gebru (2018)
Thing ethnography of AI systems	Collect data from the perspective an object in your system; use it make sense of the situation around the device; what does it experience that you did not know? Who does it interact with? Giaccardi et al. (2016); Murray-Rust et al. (2019)
Conversations with AI	Choose one of the AI-powered objects in your interaction. One team member plays the character of the object, and others carry out an interview with that object. Nicenboim et al. (2020); Reddy et al. (2021)
Metaphor shifts	Think about the metaphors use in describe your system, then think about designing purely for the metaphor; change metaphors and try again Murray-Rust et al. (2022b); Lockton et al. (2019); Alves-Oliveira et al. (2021)
Roleplaying AInetworks	Play an object, system or human role and collaboratively act out the interaction. Discover new actors, negotiations, relationships and interaction details Pschetz et al. (2019); Reddy et al. (2020)
Resisting/subverting AI	Recognise and act out moments in the interaction where someone might subvert the interaction; look for design opportunities Lupetti et al. (2020); DiSalvo (2015)
Meaningful human control	Brainstorm places where the interaction might go wrong; rather than trying to fix it, figure out who or what is responsible and how they might be supported in setting their moral boundaries. Siebert et al. (2022)

The references given here are for the theoretical context and inspiration for the methods, and do not match exactly those given on the exercise cards which prioritise developing student understanding

Fig. 3 Example of the exercises, showing (1) title, (2) expected time (3) suitable project types (4) process (5) custom illustration (6) background, (7) references and example projects (full set of exercises in Supplementary Material)

ITD

2

TIME

45 minutes

3

SUITABLE FOR

Any project

1

Metaphor Shifts

4

PROCESS

Think about metaphors to describe your AI and draw them (10 min)

Look at your system. What metaphor would you use to describe the AI? Is it a friend, a judge, a pet, a ghost? Make a simple drawing to represent each metaphor.

5

Choose one and reflect on how you would design systems based on this metaphor (10 min)

What does it imply? e.g. a pet needs feeding, a ghost is creepy. Can you use the metaphor to emphasise different parts of your design?

Now think of a completely different metaphor (15 min)

For example, what if your pet was instead an alien observer? What if your radio was a lantern shining light, or your snowboard was a magic wand? What new implications would come up?

Now try some more! (10 min)

It could be a mixture of sensible and surreal.

6

BACKGROUND

The metaphors that we use change the way that we build with technology. Structural metaphors link concepts together – 'demolishing an argument', spatial metaphors provide orientation – feeling 'up', ontological metaphors make concepts real – 'inflation is taking its toll'. The practices of machine learning and artificial intelligence draw heavily on metaphors: the neural net is an analogy between a collection of linear algebra and human thinking, or the idea of 'learning' as a shorthand for the operation of a backpropagation algorithm over those networks. These shape the way we think – 'training' a model has a very different feeling to 'fitting' it to a collection of data. New metaphors give us new ways to think.

7

REFERENCES & EXAMPLES

- Metaphors for designers working with AI, Dave Murray-Rust, Johanna Nicenboim, Dan Lockton, DRS2022, <https://shorturl.at/azA16>
- Lockton et al. (2019). New metaphors: A workshop method for generating ideas and reframing problems in design and beyond. <https://doi.org/10.1145/3325480.3326570>
- Dove, G., & Fayard, A. L. (2020, April). Monsters, metaphors, and machine learning. <https://doi.org/10.1145/3313831.3376275>
- Alves-Oliveira et al. 2021. Collection of Metaphors for Human-Robot Interaction. <https://doi.org/10.1145/3461778.3462060>

wider implications of AI, as well as across developing fluency and supporting critique (Fig. 1, description in Table 1, details in Supplementary Material). Some of the methods were pre-existing explorations, some had been used extensively and some were adaptations of existing techniques to fit the AI context or the autonomous format. There was a strong focus on activities that could be performed relatively simply by students, that were experiential, and that would work across a range of topics and levels of technical accomplishment. Each exercise was intended for application to an existing project, i.e. not a brainstorming or early ideation tool, but a way to develop existing work.

3.3 Execution and data collection

Towards the start of the third stage, after feedback from the second exhibition (Fig. 2), a half-day workshop was organised in which all student teams were introduced to the 9 design exercises with the aim of refining their project. This timing was chosen primarily for educational reasons—the methods here were designed to help develop and refine existing ideas, rather than generate new ones, so we waited until the final stage of the course. This timing was the subject of some discussion—see Sect. 4.1.2

This half-day workshop allowed each group to execute one or two of the design exercises for their design project. Output of the half-day workshop was captured on A3 templates, including a questionnaire with some first prompts on the usefulness and effectiveness of the exercises applied. The half-day workshop was setup to be executed autonomously by the student teams, selecting the exercises themselves, with lecturers present to observe and assist when necessary. All output materials of the design exercises during the workshops were collected afterwards.

Two weeks after the workshop we invited each team to select a representative to take part in a one-to-one semi-structured interview to discuss their experience and the effect it had on their project. To minimise educational disruption at a busy time and limit the possibility of coercion, we did not attempt to get full coverage, but allowed self selection by the students, in return for a €20 contribution to the teams prototyping budget. This led to 12 out of 28 teams participating in the interview. We interviewed all coaches the week before the end of the course ($n = 7$) to see what effect they had perceived on the students work. Interview questions and structure can be found in the Supplementary Material.

3.4 Analysis and evaluation

The interviews were audio recorded and both the interviews as well as the output materials from the design exercises were transcribed and analysed by a team of seven

researchers. We inductively coded both the written materials and interviews with students and interviews with coaches. We conducted collaborative thematic analysis: the coding team collectively familiarised themselves with the data and defined a shared coding scheme. At least two members of the team coded each of the transcribed materials using this scheme. Finally, coded materials were collectively discussed to synthesise insights into key themes, framed by the three levels of engagement with AI discussed earlier.

4 Findings

Our findings are structured in two parts which build on both the A3 worksheets ($n = 28$) and the student and coach interviews ($n = 12$, $n = 7$, respectively). While the analysis of the A3 sheets revealed recurring topics and common themes, the interviews revealed in-depth insights about what the students took from the exercises. The first part (Sect. 4.1) covers the execution of the methods: which ones were chosen and how they were perceived and valued by the students. The second part (Sect. 4.2) describes the links made to AI and machine learning at the interactional and relational levels as well as wider implications. In all cases, comments from student interviews are marked as [P] < [ID] > and those from coaches as [C] < [ID] >; extra context about the project that the quote relates to is given in square brackets. The students who participated in interviews were working on projects around: comfort and behavioural encouragement while driving, as well as behaviour modelling and matchmaking (for Ford); collection of data while surfing and intelligent ski clothes (for O'Neill); smart objects and energy manifestation in hotel rooms (for Citizen M Hotels); photography for reconstructive surgery (for Erasmus Medical Centre) and speculating on spirituality and life coaching with AI (for the DCODE project).

4.1 Method execution

To build context about the way the exercises were carried out, we give a quantitative summary of students' opinions of their project before the workshop, and their evaluation of the clarity of the exercises. We then look more qualitatively at two themes: students' sense of the relevance and overall evaluation of the methods and an analysis of the ways in which they found the methods useful. Table 2 summarises the number of times each method was and provides key quotes for their use in these four areas: concept development, detailing interactions, understanding AI and supporting reflection.

Table 2 Name, key references and description of each method, along with the number of groups who used it and the average score for clarity

Method	Clarity (-3..3)	Count	Concept Development	Detailing Interactions	Understanding AI	Reflection
Uncertain Inter- actions	1.1	●● ●● ●● ●	●●●●● “make it more con- crete” [P24], “a screen to show why...” [P6]	●●●●●●● “interaction in detail” [P7], “pre- pared for difficult con- versations” [P24]	●●● “better pre- pared for [] error interac- tions” [P16]	●●● “reconsider intentions” [P16]
Be the ML	2	●	● “gave us an overview” [P24]		● “listening to human emo- tions is hard” [P24]	
Poor Datasets	not used	not used				
Thing Ethno- graphy of AI Systems	2	●●	●● “the journey of the object outside the interaction” [P22]	●● “react to being held” [P22]	● “[camera] viewpoint was not suitable” [P22]	
Conver- sations with AI	2.2	●● ●● ●	●●● “takes state of mind into account” [P2], “give AI a purpose” [P3]	●●● move from passive to active [P10], feel-good to good [P2]	●● “AI cannot have the same expe- rience as you” [P3], “sensitizer” [P2]	● parts are separated, importance of roles (p10)
Metaphor Shifts	1.8	●● ●● ●● ●● ●● ●●	●●●●●●●● ●●● “role and charac- ter” [P16], “different perspectives” [P7]	●●●●●●●● ●●● “find a way to initiate nego- tiation” [P3], “gets a char- acter” [P23]	●●● “AI needs to be specific” [P18], “AI as a protective agent” [P3]	●●●●●● “how that intention has manifested itself into the concept” [P16]
Roleplaying AI Networks	1.2	●● ●● ●	●●●●● “new agents and interac- tions” [P12], “better clar- ity” [P13]	●●●●● “draw out interactions” [P13], “elab- orate on context” [P27]	● “understand- ing of specific [technical] elements” [P13]	none reported
Resisting/ Subvert- ing AI	1.0	●● ●	●●● “explor- ing new ideas” [P18], “chances to be misused” [P8]	●●● “things we knew but never wrote down” [P28], “criticality” [P8]	● adversarial sensing (p8)	● “think criti- cally” [P8]
Meaningful Human Control	1.0	●● ●● ●● ●	●●●●● “making our design more responsible” [P19], “In- troduce [...] resistance” [P2]	●●●●● “development of consent” [P25], “part of interactions we haven’t planned” [P19]	●●● “understand the richness” [P3], “mess with your moral judge- ment” [P2]	●● “consider user vulnera- bility” [P21], “bring across critical view” [P25]

Black dots (●) represent the number of groups who used the method. The final four columns address perceived usefulness of the exercise for four areas: concept development, refining interactions, understanding of AI and reflection. White dots (○) count the number of groups who mentioned it as being helpful for the area, with representative direct quotes in italics, summaries in normal font

4.1.1 Quantitative self-assessment and clarity of the methods

Analysing the worksheets, we looked into how many times each method was used as well as the perceived clarity of the instructions (Table 2). 20 of 28 groups carried out two exercises, with the remaining 8 carrying out only one. Counting responses to questions about their projects where a value greater than 0 was given (Fig. 4), 16 groups (0.57) felt their project critically investigated technology; 12 (0.43) were solving real-world problems; 15 (0.54) made use of AI qualities; 22 (0.79) engaged with complex relationships and 16 (0.57) intended to consider the wider implications of their work.

All of the methods were rated as clear (Likert scale $-3 = \text{"very unclear"}$, $3 = \text{"very clear"}$, $m = 1.5$, $sd = 1.0$, $m_{min} = 1.0$), with only a single instance being rated negatively. This indicates that students felt they understood the purpose and structure of each method.

4.1.2 Relevance and situation in the course

Most groups chose exercises to address what they considered unexplored in their projects, or in some cases, even limitations of their concepts. For instance, Metaphor Shifts was picked to “look for something else that better describes [their project]” [p1] or “finding a nice metaphor [to make] interaction with the AI more empathetic to the user” [p15]. Roleplaying AI Networks offered the hope to be “really precise and defined in the personality that we were gonna give the AI [that was deciding people’s futures]” [p27], and Uncertain Interactions was chosen to help “map out all the responses and interactions that we were not considering before” [p19] in a multi-object interaction. Other groups saw the exercises as a more general way to “check if we were in the right direction” [p1], “plan ahead like the possible problems” [p6], “have a discussion point [...] instead of just

everybody thinking in different directions” [p15] or more radically “just start again and we go somewhere else” [p6]. Some methods were explicitly avoided as the groups felt they had experienced the methods before, e.g. “role playing” [p19], or because they didn’t fully understand what a method entailed [p16]. Although the students reported that the methods were relatively clear, they suggested that the individual differences might have impacted to some level how students interpreted the exercises: “we’re all from different cultures. So, we all interpreted some questions differently” [p1].

There was a common response that this activity would have been more useful earlier in the course [p2,p5,p16,p19,p27, ...] and it could have helped generating more prototype ideas [p23]. Part of this was due to the sense that the activities felt like “an ideation—like an inspiration activity” [p2]. As students saw the exercises as tools for divergent thinking, they would have liked to use them for ideation in close connection to the prototyping experimentations in the first period of the course [p5]. Others were concerned that the moment when they did the exercises was their time to “optimise the prototype for the exhibition” [p16] and wanted to spend all of their time in making. In contrast, feedback from coaches on the timing was more positive: “To have sort of a zoomed out exercise at that point is I think a very powerful thing to do.[...] if you don’t know where you’re heading, then all these things, I don’t think they will help you. [...] So I wouldn’t move it.” [CIA]. Several students echoed this perspective emphasising how it helped their process, e.g. “because we were kind of stuck with our idea in general” [p7], that it helped “think of more details” [p20] around a developed idea.

Overall, there was a positive attitude towards the activities, even from groups that were initially suspicious: “we were quite surprised because we were thinking ‘Ohh workshop again. [...] What’s going [to come] out of it?’ [...] And then in the end there was actually some things that really

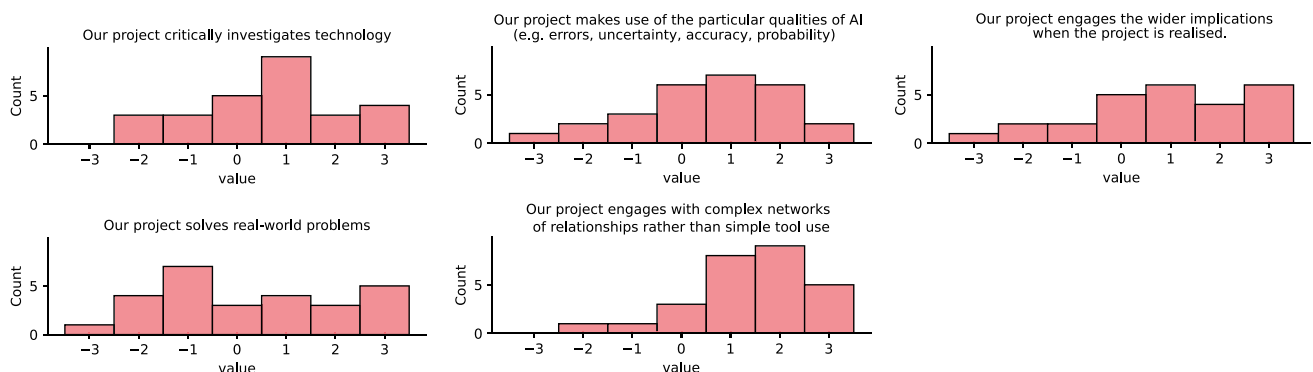


Fig. 4 Student response (per group, $n = 28$) to questions about their orientation and their project scope. Answers are on Likert scales from ‘Not at all’ (-3) to ‘It’s the core of our project’ ($+3$)

helped us.” [P1]. Some negative responses (4) revealed student’s concerns on carrying out the exercise properly [P5], or spending too long on one interaction [P19]. Some (2), had a hard time finding usefulness in the experience [P15,P16], as they were already familiar with the methods, as “*wouldn’t say it brought me a new understanding because the metaphor is something we [already] had*” [P15].

4.1.3 Perceived utility of the exercises

Many students saw the exercises as a form of “*ideation, like an inspiration activity*” [P2], “*kind of a brainstorm*” [P7] that can help “*to get a better idea*” [P7]. Several groups noted that they came up with different and more interesting ideas [P5] and that they “*could use this new inspirations*” [P23]. The methods were seen as useful for sharpening projects and defining practical next steps, such as planning for when things went wrong, or checklists of common concerns. Metaphor Shifts was particularly generative of new ideas [P5,P16]. Beyond this, the methods were seen to help in the following four areas:

Conceptual development: The methods supported articulation and “*helped us get our story right, like the overall purpose of the concepts*” [P2], to develop “*a better detailed new metaphor*” [P5] and to “*make [a] choice in what we wanted*” [P7]. The benefit of gaining more conceptual clarity was also mentioned by some coaches [CLU]. Uncertain Interactions was seen as useful for mapping out the edges of a concept, so that students could easily get into details and next steps [P6]. The methods also helped grounding ideas, asking whether the concept “*can also work on the AI or do we need some future technologies that are not there yet to make it real*” [P1]. In some specific instances, the exercise helped to “*start thinking about time*” [P15], or to “*find new ways of taking the same idea and spreading it*” [P5]. The exercises helped to get an overview of things that students should think about [P1], making implications concrete and graspable, in a way that is “*so in your face that you don’t even think about the fact that it will be in the future*” [P7].

Refining interactions: Many groups came out with a more refined idea about how their conceptual interaction should play out, as the exercise “*asks you to go into parts that maybe you don’t want to explore*” [P6] and make projects more well rounded. The activities also helped students define interaction contexts better. Groups felt invited to “*draw [AI] already in the context*” [P15], and to “*think about the interaction with some of the objects in the [interactive hotel room] scenario*” [P19]. Refinements also pushed them to account for the potential meaningfulness of the projects, to “*clarify intentions*” [P20] and anticipate outcomes, e.g. “*what happens if the user doesn’t understand what [the smart objects are] talking about*” [P19]. The experiential nature of the exercises helped to “*translate something*

abstract as “being challenged” or “supporting” [good behaviour while driving] to something actually tangible” [P2], to think into the “*aesthetic experience*” [P15] of AI where the “*metaphor [of ritual cleansing for data collection] helped to think about materials as well*” [P15].

Reflection: The workshops were seen as a moment of reflection, a break from the “*many layers in such a project*” [P25] to focus on particular aspects. This could be on a technical level for the groups who “*never really took the time to think about AI*” [P2] or more interactional when they “*stopped to think about this character sort of thing*” [P27]. There was developing a “*critical lens, in terms of moral responsibility*” [P25] and “*seeing how important this is, to acknowledge the mistakes, to be trustworthy*” [P6]. Beyond the initial designerly sense of responsibility, they engaged with broader factors contributing to “*moral responsibility for an AI system [that encouraged spirituality]*” [P25]. Overall, the moment for reflection was seen positively, developing aspects of their work that were not thought through, and a sense that “*confidence comes once you [...] manage the critical points*” [P6] of the interactions.

Understanding AI: The workshop improved the confidence of students about working with AI, as “*before the course it was just like ‘I don’t know how to use an algorithm to do something cool’ [...] and this makes it kind of [makes] everything just specific in one one workshop*” [P2]. This was often not based on a deeper technical understanding of algorithmic operation, but on a thinking about how the AI would relate to things around it. Some groups ended up “*actually using more AI because of this [workshop]*” [P7], with confidence coming from “*now that we know what’s going wrong, and we know how to respond to that*” [P6].

4.2 Key themes for engaging with AI

Now, we discuss findings in relation to broader theoretical developments in HCI, according to the three levels we have identified earlier: interactional affordances, relational questions and wider implications. An overview of these findings can be seen in Fig. 5

4.2.1 AI interactional affordances

Students found that the workshops illustrated that “*there are actually a lot of possibilities with AI*” [P7], beyond the tutorials at the start of the course, and that working through the experiences left them with a “*whole list of things that [AI] could say or do*” [P2]. They already had some experience with particular topics, but this opened up a greater sense of how these possibilities could be deployed in relation to their work. This did not always change the concept of the interaction, but did provide a confidence that many interactional designs could potentially be realised.

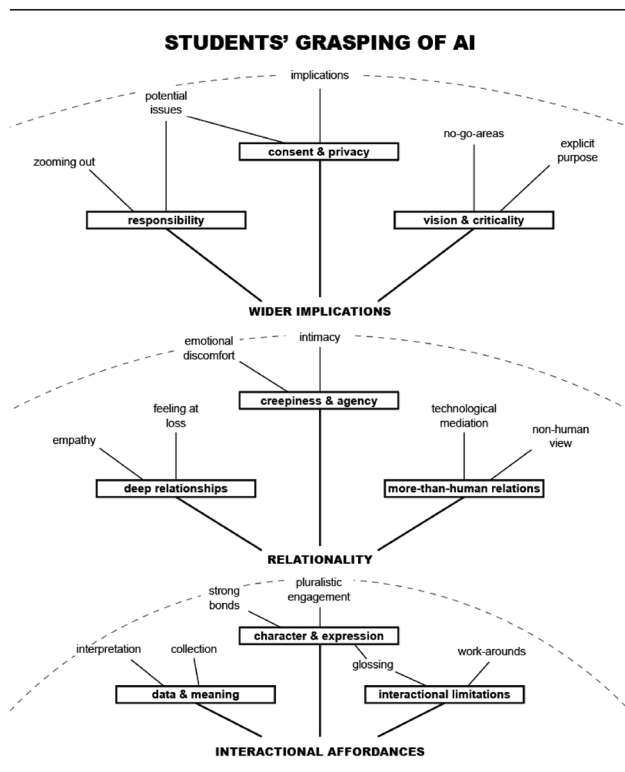


Fig. 5 Conceptual map of students' reflections on the benefits of the methods grouped across the three levels of AI engagement: interactional affordances, relationality and wider implications

Data and meaning: Role playing helped with sensitising the students to the role of data in AI driven systems, questioning “where is the AI getting the information?” [P27] both generally and through very detailed questions of “where we’re gonna put [the camera that understands human–vehicle interactions]” [P5]. The coaches noticed the attention to physical detail as well, seeing development of a “way of bringing the data and looking at it and experiencing it” [CIA]. There were moves to think about how to work with people in wheelchairs, and what it would mean to “recognise these things and build the dataset” [CLU], as well as the broader question of how “how [the collected data] can be meaningful for you as a person” [CIA].

Character and expression: The experiential nature of the exercises was, unsurprisingly, suited to engaging with designerly questions of the character and expression of the autonomous parts of the system. Students notice the possibility that they could be “really precise and defined in the personality that we were gonna give the AI” [P27], questioning default assumptions about how the system might respond. With conversational agents, it was noted that “there is a lot of space between the yes or no” [P19], but also that working probabilistically could smooth out interactional challenges, so that humans “don’t have to become machines ourselves”

[P6]. The possibility arose to create pluralistic engagements that gave “different answers based on different characters and based on different situations [for patients undergoing reconstructive surgery]” [P23]. This opened the possibility of making stronger bonds with users, and working on an emotional level, which we will return to in the next session.

Interactional limitations: In general, the coaches were more sensitive than the students to the potential limitations of technology, for example noticing when “the way they acted out looks good on screen but it doesn’t reflect the deeper issues with understanding [...] Whereas if you use a conversational AI model I think you will run into a lot of problems that are hard to act out” [CGI]. It was clear to them that some of the enactments would require sophisticated behaviour that could easily be glossed over with WoZ techniques, and they questioned whether the exercises could also point to these moments of glossing, or help notice points of complexity. For the more technically realised groups, the coaches noticed students working around limitations of the technology, where “it was not very good at detecting facial expressions, but you made a hand gesture” [CGI] that conveys emotion purposefully, leading to a rethinking of the interaction schema.

4.2.2 AI relationality

Students felt that “[t]here are so many layers in such a project, where you are constantly building” [P25] and noted that the workshops took them into some of the complex, multi-layered aspects of working with interactive AI systems.

Deeper relationships: Following the theme of character above, the workshops prompted students to think about the ways that humans related to the things being designed, giving an impetus to “think more in an empathetic way” [P15] about the end users and what AI mediation would “mean for a human to human relationships” [P25]. Roleplaying the situation with the device helped to look across some of the other people around the interaction, for example working with an system that was helping to take medical photos for reconstructive surgery and seeing “the relationships between the AI [and] doctor, assistant to friends or to your family members” [P23]. This was partly driven by a sense that the AI systems could interact in increasingly human-like ways, with metaphors like “a friend in your car” [P5], or a pet. There was a move to look at some of the longer-term relationships formed and the bonds that people made with AI systems. Students developed increasingly anthropic concerns from whether “people feel at a loss after they need to give [their smart mirror] back” [P23] at the end of a process, to questions of developing care and love relations to the objects.

Creepiness and agency: Interestingly, some of the more than human metaphors helped students to about when

agency was troubling, and were “*open to more scenarios that we didn’t see*” [P20]. Manifesting home energy use using a metaphor of ‘fireflies’ caused a concern that it “*will follow you through your room as a dog follows you. This might be kinda creepy. So what if they [users] don’t want to be followed?*” [P20]. The potential intimacy of relations with a vehicle raised concerns about “*how intimate your interaction with your decentralised car be*” [P5], and how “*if you’re driving and you’re stressed and you somehow just get like this random unexpected hug from your car*” [P5] it would cause emotional discomfort. Even when autonomous behaviour was not emotionally invasive, there were concerns that “*sometimes the [smart hotel room] objects want to speak for themselves but at the same time you don’t want to scare the human that is the guest in this room*” [P19].

More-than-human relations: Going beyond metaphors of caring for cars as one might care for a dog, coaches noticed that students would “*use design as a medium to amplify the voice of nature*” [CIo] or “*activate [...] energy consumption in a different way than just a tool*” [CLU] in their AI mediated interactions, making a shift to both non-human perspectives and the idea of technological mediation rather than tools for particular outcomes. The students looked into new relationships that might emerge, e.g. designing “*clothes to learn from every person that wears them to [and] grow its own personality*” [P16]. The coaches noticed the roleplaying aspect of the exercises prompted critical reflection into the scenarios and relationships at hand, including noticing “*that the setting that they were imagining and the role of the AI within that setting was not a very good fit*” [CGi]. Students found the practice useful for articulating what their vision for the future of human–AI relationships at individual and societal levels ought to be, including questions of governance and democracy.

4.2.3 AI and wider implications

Responsibility: While methods targeted at interrogating control (Meaningful Human Control) explored agency and control, other methods (Metaphor Shifts) still gave space for these questions to arise. Students reflected on “*considering moral responsibility for an AI system*” [P25] within the creation process; and the coaches noticed that the workshops provided “*a way to create distance and look at the project from a different perspective*” [CIA], to re-evaluate the project beyond the immediate concerns of development, with a sense that it was the designer’s responsibility to make sure that purposes and potential issues were clear upfront. Some students found that the workshops made the idea that people might misuse their system concrete, so for a friendly car system they “*gave ourselves some guidance for the next steps, [not] for concepts [...], but more like OK, this is now a checklist that we need to put next through concept every time*

to make sure we think about this” [P1]. Responsibility often came through thinking through what might go wrong, with evidence of ‘zooming out’ through the exercises, to think about what would happen if these systems were widespread, and their failure modes constantly present for users.

Consent and privacy: Several students mentioned issues around consent; while some felt this was a core part of their existing work, others found that discussion around the workshops was what they needed really understand the implications, and “*a solution for something that [is] difficult to think about*” [P1]. Groups managed to “*dig deeper in that space*” [CLU] and better manifest the issues that they were already dealing with, and in some cases this meant that “*[consent] was actually a very explicit part of their final concept and that was not at their departure, I think, was driven in part by going a bit more speculative than they were imagining at first*” [CGi].

Vision and criticality: A common point from the students was that these workshops helped to think beyond the initial concerns of prototyping and into the multi-layered nature of the projects, not just the around AI responsibility but that “*it asks you to go into parts that maybe you don’t want to explore*” [P6] and rethink the purpose and shape of the project itself. Coaches were mixed about whether they saw changes in the level of critical thinking around the workshop, with some noticing no change, some a progression, and some seeing a strong difference where critical thought was brought in. Some of these were trade-offs: “*They became more critical. They were focussing more on the experience, but I’m not sure they were more engaged with the AI*” [CIo]. However, others noticed engaged with the human–AI relationships, questions of datasets and the role of the project as critique, and “*really thought about it, how you negotiate with the machine and how much freedom you should have and how much agency you just have*” [CMA].

5 Discussion

In this discussion, we address some potentials for developing the exercises, and reflect on our initial research questions. We discuss how the AI exercises address the current methodological gaps and, more broadly, how this work contributes to a larger programme around design, HCI and AI, nurture a distinctively designerly AI culture.

5.1 Effectiveness and future work

The exercises were seen as effective overall, although they could further be improved through use, observation and iteration. They produced thoughtful, socially engaged responses, but to a large extent remained far from the rapid and technically grounded results generated at the beginning

of the course, when students were provided with tutorials focused on learning a particular AI technologies. As an example, despite the deep technical grounding of *Uncertain Interactions*, most student responses did not get deep into the specifics of model output and how to make use of it. Future versions of the exercises could look to bridge this gap, as could their use in more technical contexts, where models were really being trained and deployed. There could also be support to help students to decide which concerns to prioritise—for example, worries about people falling in love with their AI devices might not be the key problematics of the technology as created. While this prioritisation is arguably a part of general design practice, having concrete examples to contextualise the discoveries would be helpful. Practically, most students were relying on pre-built models and ‘Wizard of Oz’ setups (Browne 2019; Dahlbäck et al. 1993) that used human action to simulate complex behaviour. This limited the utility of data-driven exercises (e.g. *Poor Datasets*) and fed into a focus on the anthropomorphic possibilities of AI. This also led to less engagement with the possibilities of new forms of human–machine interaction than we might have hoped for.

5.1.1 Timing and situation

The time that the students had to execute the exercises was short, which may have limited the potential for deep reflection and thoughtful practice. While the students still had access to the methods, few groups chose to make use of them, so there is space to explore more prolonged engagement. The positioning in the course was somewhat contentious, with many students feeling the methods had been introduced too late (Sect. 4.1.2)—this is coupled to their assumption that the methods were there for concept development and ideation. However, the overall feeling from the coaches was that the timing was sensible—it provided a way to zoom out around existing concepts and add richness. Part of this divergence of opinion is part and parcel of process based education—there are often different views from within the process than outside it. However, it does point to the need for a stronger sense of what one can expect from the methods, and an indication of when and how they could be productively deployed.

5.1.2 Choice and range of methods

This initial set of exercises was based on a particular set of theoretical ideas; it is clear that other theories and concepts could prompt additional methods, and other methods could be derived from the theories used. There is certainly no shortage of candidates, whether agential cuts (Shotter 2014) provide techniques to divide up complex systems and consider multiple boundaries through more or less embodied

encounters (Vagg 2022), or ideas of cyborg intentionality (Verbeek 2008) lead us to enact pairings with composite possibilities (Rapp 2021), introspection provides a lens to think about relations between AI and lived experience (Brand et al. 2021). Methods with a clear technical genesis would offer immediate experiences that are deeply embedded in and shaped by the technology, for example deliberately mis-using vision algorithms (van der Burg et al. 2022) or using computer vision as a site of enquiry (Malsattar et al. 2019). We see this as the start of a collection of ways to engage in this area, which will grow over time. Additional exercises might emphasise different parts of the design process and different modalities of experience as well as introducing new theories or grappling with particular qualities of AI.

5.1.3 Applicability

In terms of subject matter, the exercises were applicable to a range of projects across autonomous cars, robots, Internet of Things, hospitality and so on. They also helped with a range of issues, from shaping overall concepts to detailing important parts of the interaction. The application here was somewhat particular: the middle stages of an exploratory, creative prototyping brief. We would expect that the methods can be used in other processes and different levels of technical fidelity. In fact, several of the methods, such as roleplaying AI networks and *Thing Ethnography* of AI systems, are likely to give better results as the project is more developed and the context is stronger. Others, such as *Poor Datasets* are likely to be more useful with a developed technical implementation, while *Uncertain Interactions* could help with ways to create interfaces around probabilistic models in deployment.

5.2 RQ1: Conceptualising and prototyping practices with AI

The exercises illustrated some of the issues that students have when carrying out prototyping and conceptualisation with AI: the need to deal with uncertainty, the possibilities of more human-like interactions but less clearly defined capabilities, the need to hold multiple levels together. This clearly asks a lot from designers, especially in this case, where many of them did not have strong electronics and coding skills before the course. The experiential (YHement et al. 2022a) and enacted (Elsden et al. 2017) aspects of the workshops were helpful to navigate this terrain, as the subjects of discussion could be played out in the group, adding to the sense of tangibility and refining how interactions should unfold. The interactional focus of this work makes it distinct from ideation tools such as *AIxDesigns* ideation cards (AIxDesign 2022) which focus on conceptual innovation or work on developing user experiences (Subramonyam et al. 2021b, a) which makes the interface the

primary subject of design. In line with open ended, critical and speculative prototyping methods (Malsattar et al. 2019; van der Burg et al. 2022; Nicenboim et al. 2020) the exercises took the students into the relational and interactional possibilities of AI.

From the feedback, it was important to give students exercises that were concrete enough that they could follow the steps. Several of the exercises were close to relatively standard design practices—Uncertain Interactions drew on the creation of state diagrams as a design articulation tool and the idea of acting out interactions as a form of prototyping is well established (Van Der Helm and Stappers 2020). However, they were adapted to bring AI qualities into the familiar interaction design practices, emphasising aspects like uncertainty, interface capabilities, distributed responsibilities and so on. It is clear that for some of the students, the simple forms of the exercises would have been enough—simply asking ‘what might go wrong’ and drawing a state machine to deal with it, rather than getting into the idea that machine learning systems produce probabilistic outputs produced useful results. None of the students chose to work directly with datasets; this may be a feature of their projects, as there was not much training and learning happening, or simply a lack of attraction to the particular exercise.

There was a tendency with many of the groups to drift into anthropomorphisation, to imagine relations as overly human (Marenko and van Allen 2016), and be diffuse about the capabilities of the technology. This relates to thinking into some of the particular AI characteristics that we will discuss in the next section, it is clear that prototyping will start to take different forms. The evolution of prompt engineering as a discipline (Liu and Chilton 2022) and the potential to generate working systems from prompts (e.g. Aptly aptly 2022) indicates that new forms of prototyping are emerging. Here, the constraints are less well defined than working with code on Arduino, but no less present - training a TeachableMachine (Carney et al. 2020) to detect a gesture has just as many concerns as using the electronic gesture sensor built into the Arduino BLE Sense the students were using, but the failure modes play out differently, and a different set of prototyping practices are brought to bear. The multiple viewpoints contained in the exercises here—people, things, datastreams, algorithms, networks—help to tease out the parts of interactions to prototype. Enacting these possibilities makes it easy to fall into broad, fuzzy, anthropomorphic thinking about what systems might do; the challenge for developing new forms of prototyping is to temper this with a grounding in the capacities of the systems being designed, and to engage with the human-like affordances of technology, without missing the new machinic possibilities. The exercises here helped students to clarify their concepts, move forward with their prototyping, and develop ideas about the responsibilities of creating AI systems, while maintaining

designerly concerns of materials, aesthetics, function, fit to context and engagements with multiple actors.

5.3 RQ2: Grasping interactional, relational and contextual qualities of AI

Our analysis of the student responses in relation to the current paradigm shifts in HCI, shows that interactional, relational and contextual qualities of AI could be important elements of design and AI educational programmes. To unpack this, we look at our findings in relation with wider notions of agency, human–machine relations, and understandings of AI.

5.3.1 Agency

As noted above, AI is a tricky term, but a lot is contained in the ideas of agency which it can develop, in particular around ‘non-humansque agencies’ (Hildebrandt 2020). Some of these agencies were clear prompts for our work: a failure to recognise certain kinds of people as being human (Buolamwini and Gebru 2018) shapes the inter-agencies between vision systems and people; the ability to make decisions rapidly and constantly gives a sense of autonomy, but one which differs both in character and meaningfulness from the one of humans. Several of the exercises were aimed at interrogating these questions: ‘Be the AI’ prompted a reflection on exactly what the machines were doing, the interviewing and roleplaying exercises asked the participants to feel into what the agential possibilities were, and the more conceptual exercises questioned what agencies and responsibilities humans had around the systems. This helped participants to think about “*what are the actual choices that we are gonna make or what part of the interaction are we gonna do ourselves and what part is the machine gonna do?*” [p16]—a key part of getting past the myths about AI capabilities (Natale and Ballatore 2020).

Much of the thinking reported around issues of agency had to do with “*how it’s gonna be alive for people*” [p20]—the clear, animate, characterful side of agency. This surely has roots in some of the roleplaying methods. The shape of the collective roleplaying exercises was informed by an increased emphasis on co-performance, as humans were brought in to act out what the smart technologies would do, and notice possibilities for more shared agency, whether co-learning with the AI or finding ways for objects to speak for themselves.

5.3.2 Human–machine relations

Some of the exercises prompted students to position AI in relation to humans and non-humans. Thing Ethnography of AI systems instructed the students to map the ecosystem of

the thing and its touchpoints, and reflect on who and what interacts with the concept. In *Conversations with AI*, the students were asked to enact an AI agent. *Metaphor Shifts* asked students to design systems based on a particular metaphor and then compare it to others. These exercises highlighted the relations of humans and non-humans within AI systems (Coskun et al. 2022; Nicenboim et al. 2020). They did that by decentering the designers perspective, to consider more actors and interactions that go beyond one user and one device (Verbeek 2020). They also invited students to relate to AI not only as a tool, but as a social agent that shapes people's lives. In students' prototypes, intelligence, as well as responsibility, were not seen as a properties of machines alone, but shared between humans and artificial partners. Similarly, uncertainty and unpredictability were 'collaboratively curated' to 'imagine forms of digital interaction' (Marenko and van Allen 2016).

The findings show that the exercises helped students to expand their concepts to account for the ecologies around them, especially when their projects were centred on particular embodiments, for example extending from a plant pot to a community of plants, and looking at the plant–plant relations as well as the plant–human ones. The exercises also helped them acknowledge other people beyond the immediate users, thinking into how would they relate to the system, and what are the responsibilities that the user, system and designer have towards them. Furthermore, thinking of their concepts in relation to humans and non-humans, created awareness within the students on the human labour that is implicated in sustaining AI systems (Sinders and Ahmad 2021). The kind of metaphorical social relationships the AI had with others ultimately influenced the designs: when the system was cast as a friend, it was seen, designed and conceptualised differently from when it was cast as a pet.

5.3.3 From explanations to understandings of AI

One of the current challenges in the design of AI systems is how to support people in understanding them, especially when used to make autonomous decisions or create knowledge. AI explainability is especially challenging when based on deep learning models, given that some of the paths that AI systems use to give recommendations are not interpretable (Ehsan and Riedl 2020), and the source of many generative outputs is complex (e.g. Kovaleva et al. 2019). While understanding ML in its technical sense is important, recent approaches in the explainability of AI have pointed at other ways of understandings which are not based on technical explanations and instead, promote experimentation, challenging boundaries, or promoting respect (Nicenboim et al. 2022; YHemment et al. 2022a; Seymour et al. 2022). The findings expand the agenda of Explainability of AI by illustrating and unpacking particular design engagements with

AI that go beyond mastering ML technical capabilities. This points at particular aspects that are important in the kind of understandings that designers might need to gain of AI. Some particular aspects that helped designers understand AI are exercises that could prompt reflection into the affordances, relations and wider implications that those systems might have. Those engagements were not based on learning how to code ML models, but on experimenting with changing perspectives, provoking failures, enacting behaviours, and drawing schemas. These tactics could become part of a new agenda for supporting designers in understanding AI, especially one that is aligned with theoretical developments in HCI such as the posthuman turn (Lindgren and Holmström 2020) as well as practical developments in design (such as methods used in critical, speculative and adversarial design (DiSalvo 2015; Irani and Silberman 2014; Bozic Yams and Aranda Muñoz 2021)).

5.4 RQ3: Critical design perspectives while engaging with AI as a socio-technical system

While the exercises helped the students to develop their projects (from ideation to conceptualisation and detailing), they especially illuminated and modulated changes in the students design processes in relation to the socio-technical aspects of AI systems (Crawford 2021). The exercises supported students in reflecting on the role of AI within their concepts, in being more specific in what kind of aspects of AI are present, and developing a critical design perspective on AI around values of responsibility and agency.

Designing with AI as a socio-technical system means acknowledging that it is not only a technical domain, but also entangled with social practices, institutions and infrastructures, politics and culture. AI "is both embodied and material, made from natural resources, fuel, human labour, infrastructures, logistics, histories, and classifications" (Crawford 2021). This is not an entirely new perspective—AI has long been considered as a material practice (Agre 1997), but there is a need to consider the interaction between humans and machines as part of broader societal contexts, and the broader discursive settings in which AI is socially constructed as a phenomenon with related hopes and fears (Lindgren and Holmström 2020).

From the findings, it seems the exercises provided a space for students to go beyond the immediate concerns of a rapid prototyping session and engage in reflective practices that can position AI within its broader societal contexts. It is clear from many of the responses that the workshop carried out here provided a moment to reflect. Some of this can be simply ascribed to the sole fact of having the intervention—a space that prompted further thought. However, some of the exercises were more specific triggers for reflective

engagement, with Meaningful Human Control and Resisting/Subverting AI asking students to explicitly consider critical perspectives. There was evidence of ‘reflection-in-action’ (Yanow and Tsoukas 2009), moments where in the midst of carrying out experiential exercises around the prototypes they ‘ma[de] previously implicit assumptions about the work explicit’ (Wegener et al. 2019). Where it had previously seemed that an insect swarm would give a warm sense of companionship, looking at the technological sense of surveillance and following revealed a darker possibility for the user; the idea that a supportive hug came from technical rather than human agency was found on reflection to be disturbing.

Some of the exercises prompted a sense of ‘zooming out’ (Nicolini 2009), to consider wider networks of things and people, and this zooming out was part of the students move towards more empathic design. This touched on the temporal (Pschetz and Bastian 2018) aspects of their design, as they thought not just about interactional moments, but the slower unfolding of relationships over time. Overall, there were many moves to develop a sense of criticality within their design: the AI oriented methods created viewpoints for considering the role of technology and its possible overreaches from pragmatic and whimsical perspective, in line with the ‘ongoing practical, critical, and generative acts of engagement’ (Suchman 2020) that build a responsibility for the things being designed.

5.5 Beyond education: nurturing designerly AI cultures

We started with the aim of helping students to grasp enough qualities of AI to adapt their processes and conceptualisations accordingly. The study highlighted that there are distinct ways to engage with AI that are appropriate to our setting, where the culture and practices of designers centre particular ways of working. In this section we discuss the aspects of our work that nurtures this designerly AI culture.

The use of AI technologies varies by field. If we look at AI in terms of the enabling technology and the culture that surrounds it (Caramiaux and Alaoui 2022), some of the differences and parallels with the use of AI in design become clear. There are common moves to cast AI as a creative partner (Llano et al. 2022; McCormack et al. 2020) within music, as a solution for optimisation (Noor 2017) within engineering, as a formalisation and purification of human thought (Chiusi 2020; Singh et al. 2019) in decision making organisations and so on. Within design, the places that AI might sit are being negotiated. Do we bring it into the process as a sparring partner for ideation (Simeone et al. 2022) or a source of creative inspiration (Yun et al. 2022)? Do we use it to re-understand the world through divergent practices (Malsattar et al. 2019)? Is it a new computational capacity for which we have to develop

new UX practices (Subramonyam et al. 2021b)? Or a boundary object whose politics need critique (Crawford and Paglen 2019; Lyons 2020). All of these are within the remit of design. What we are interested in accenting here is the possibility for a designerly culture around the use of AI technologies, whether in processes, outcomes or critique. Just as a shift from explanation to shared understanding (Nicenboim et al. 2022) speaks to a relational, experiential mode of engagement, the exercises here create those experiences, and give ways to pick up, tangle and hold those relations. We suggest there are three key features of the methods that support this: experientiality, pragmatism and reflection.

The experiential nature of the methods appeared to be key in bringing in different perspectives on existing work, from noticing potential implications to uncovering new actors and interrogating positive ideas of agency. In this prototyping oriented style of working, enacting and dramatising possibilities helped to grasp concepts. This was particularly relevant to working with AI systems, where the level of agency expected of the technology is high, so vitalising it makes intuitive sense.

Second, the pragmatic nature of the exercises, distilling complex ideas down to a set of steps to explore supported critical discussion. Rather than starting from the theory, students were able to develop grounded experiences and respond to them. This led to practices such as developing their own checklists for responsibility as well as rethinking interactions based on new metaphors for the relations between technologies and humans.

Finally, the exercises all point to building the skills that a reflective designer in AI might need—“*Perhaps the thing that they have in common is that they make you reconsider what your intention was and how that intention has manifested itself into the concept*” [p16]. As such, they are distinct from technical support, even technical support tailored to creative practitioners (AIXDesignComm 2020), but look to build bridges from more than human thinking (Coskun et al. 2022; Coulton and Lindley 2019; Giaccardi and Redström 2020; Nicenboim et al. 2020) towards technical practice.

By providing this kind of multiple toolbox, we contribute to shaping the emerging AI-Design culture as something distinct from the technical, scientific, artistic and socio-legal cultures that are relatively well established. Further, we believe that this practice of grasping AI can be useful beyond the classroom, a powerful and versatile support for design professionals to meaningfully engage with the development of intelligent systems.

6 Conclusions

There is a growing need for designers to engage with artificial intelligence and machine learning in their practice as it becomes integrated into the functioning of the physical

and digital systems that they design. A particular challenge here is how to carry out ideation and early stage prototyping around AI/ML, when the exploratory nature of the work makes it impossible to invest much time in detailed technical understanding of particular algorithms or systems. At the same time, the technical possibilities of emerging algorithms can exert an overly large pull on designs, artificially narrowing the solution space and drawing away from the needs and qualities of the interaction.

To develop the potential for designers to engage meaningfully in this space, working from an educational perspective, this paper introduced a series of ‘AI exercises’ informed by recent theoretical developments in third wave HCI to help students grasp AI as a socio-technical system. We developed three levels of consideration for designing AI systems: interactional affordances, relational possibilities, and the wider social implications of AI systems; and provided methods for working at each level. Through qualitative analysis of these exercises with a group of students, we build up a picture of what kind of impact the interventions had on their understanding of AI and their project development. Through the exercises, the students refined their designs and clarified their concepts, and were able to move forwards with their prototyping with a greater sense of confidence in their designs and responsibility around the process. The experiential, pragmatic aspects of the exercises helped to make theoretical ideas concrete and generative of new possibilities, while keeping a sense of materiality and interaction with humans. The space for reflection provided by the exercises helped the students to develop a wider perspective on their work within the bounds of a rapid prototyping project.

The study findings highlight the ways in which experimental design exercises could support students in understanding AI, especially considering that such understanding needs to go beyond mastering ML technical qualities. The exercises here helped illuminate and modulate changes to the students design processes in relation to the interactional, relational and contextual qualities of AI, helping students develop a reflective and critical design perspective while responding to the key theoretical developments that are discussed in the AI community within HCI. Through the discussion, we raise questions of how a socio-technical view of AI, through ideas of agency and relationality can support a designerly culture around the development of AI.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00146-023-01794-y>.

Acknowledgements We would like to thank our course collaborators for assistance in both running and analysing the course, in particular Ianus Keller, Aadjan van der Helm, Tomasz Jaskeiwicz, Dieter Vandoren, Gijs Huisman, Nazli Cila and Martin Havranek, as well as Seowoo Nam for graphic design and data collection around the methods. Thanks to the others in the Human Centred Design department who helped us to think about and frame this work, in particular Elisa

Giaccardi. This work was partly supported by the Microsoft Research PhD fellowship awarded to Iohanna Nicenboim. Finally, as an piece of educational research we would like to thank the students on the course for their hard work, thoughtfulness, creativity and boldness.

Data availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Agre P (1997) *Computation and human experience, learning in doing*. Cambridge University Press, Cambridge
- AIxDesign (2022) AI ideation cards. <https://www.aixdesign.co/shop>
- AIxDesignComm (2020) AIxDesign community. <https://www.aixdesign.co>
- Alves-Oliveira P, Lupetti ML, Luria M, Löffler D, Gamboa M, Albaugh L, Kamino W, K Ostrowski A, Puljiz D, Reynolds-Cuéllar P, Scheunemann M, Suguitan M, Lockton D (2021) Collection of Metaphors for Human-Robot Interaction. In: *Designing Interactive Systems Conference 2021*, ACM, Virtual Event USA, pp 1366–1379. <https://doi.org/10.1145/3461778.3462060>
- Amershi S, Weld D, Vorvoreanu M, Fourney A, Nushi B, Collislon P, Suh J, Iqbal S, Bennett P, Inkpen K, Teevan J, Kikin-Gil R, Horvitz E (2019) Guidelines for human-AI interaction. In: *CHI 2019*. <https://www.microsoft.com/en-us/research/publication/guidelines-for-human-ai-interaction/>
- aply (2022) Introducing aply that turns ideas into working apps. <https://appinventor.mit.edu/blogs/hal/2022/03/21/Aply>
- Auger J (2013) Speculative design: crafting the speculation. *Dig Creat* 24(1):11–35. <https://doi.org/10.1080/14626268.2013.767276>
- Auger J (2014) Living with robots: a speculative design approach. *J Hum-Robot Interact* 3(1):20–42. <https://doi.org/10.5898/JHRI.3.1.Auger>
- Barad KM (2007) *Meeting the universe halfway quantum physics and the entanglement of matter and meaning*. Duke University Press, Durham
- Benjamin JJ, Berger A, Merrill N, Pierce J (2021) Machine learning uncertainty as a design material: a post-phenomenological inquiry. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ACM, Yokohama Japan, pp 1–14. <https://doi.org/10.1145/3411764.3445481>
- Benjamin JJ, Biggs H, Berger A, Rukanskaitė J, Heidt MB, Merrill N, Pierce J, Lindley J (2023) The entoptic field camera as

- metaphor-driven research-through-design with AI technologies. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, CHI '23, pp 1–19, <https://doi.org/10.1145/3544548.3581175>
- Benjamins R, Barbado A, Sierra D (2019) Responsible AI by design in practice. [arXiv:1909.12838](https://arxiv.org/abs/1909.12838) [cs]
- Bilstrup KEK, Kaspersen MH, Assent I, Enni S, Petersen MG (2022) From demo to design in teaching machine learning. In: 2022 ACM Conference on Fairness, Accountability, and Transparency, Association for Computing Machinery, New York, NY, USA, FAccT '22, pp 2168–2178, <https://doi.org/10.1145/3531146.3534634>
- Bowler RD, Bach B, Pschetz L (2022) Exploring uncertainty in digital scheduling, and the wider implications of unrepresented temporalities in HCI. In: CHI Conference on Human Factors in Computing Systems, ACM, New Orleans LA USA, pp 1–12, <https://doi.org/10.1145/3491102.3502107>
- Bozic Yams N, Aranda Muñoz Á (2021) Poetics of future work: blending speculative design with artistic methodology. In: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, CHI EA'21, pp 1–8, <https://doi.org/10.1145/3411763.3443451>
- Brand N, Odom W, Barnett S (2021) A design inquiry into introspective AI: surfacing opportunities, issues, and paradoxes. In: Designing Interactive Systems Conference 2021, Association for Computing Machinery, New York, NY, USA, DIS'21, pp 1603–1618, <https://doi.org/10.1145/3461778.3462000>
- Browne JT (2019) Wizard of Oz prototyping for machine learning experiences. In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, CHI EA'19, pp 1–6, <https://doi.org/10.1145/3290607.3312877>
- Buolamwini J, Gebru T (2018) Gender shades: intersectional accuracy disparities in commercial gender classification. In: Conference on fairness, accountability and transparency, PMLR, pp 77–91
- Caramiaux B, Alaoui SF (2022) “Explorers of unknown planets”: practices and politics of artificial intelligence in visual arts. In: Proceedings of the ACM on Human–Computer Interaction
- Carney M, Webster B, Alvarado I, Phillips K, Howell N, Griffith J, Jongejan J, Pitaru A, Chen A (2020) Teachable machine: approachable web-based tool for exploring machine learning classification. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp 1–8
- Cavalcante Siebert L, Lupetti ML, Aizenberg E, Beckers N, Zgonnikov A, Veluwenkamp H, Abbink D, Giaccardi E, Houben GJ, Jonker CM, van den Hoven J, Forster D, Lagendijk RL (2022) Meaningful human control: actionable properties for AI system development. AI and Ethics. <https://doi.org/10.1007/s43681-022-00167-3>
- Chiusi F (2020) Automating Society Report 2020. Tech. rep., Algorithm Watch, <https://automatingsociety.algorithmwatch.org>
- Coskun A, Cila N, Nicenboim I, Frauenberger C, Wakkary R, Hassenzahl M, Mancini C, Giaccardi E, Forlano L (2022) More-than-human concepts, methodologies, and practices in HCI. In: Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, CHI EA '22, pp 1–5, <https://doi.org/10.1145/3491101.3516503>
- Coulton P, Lindley JG (2019) More-than human centred design: considering other things. *Des J* 22(4):463–481. <https://doi.org/10.1080/14606925.2019.1614320>
- Crawford K (2021) The atlas of AI: power, politics, and the planetary costs of artificial intelligence. Yale University Press. <https://doi.org/10.2307/j.ctv1ghv45t>, <https://www.jstor.org/stable/10.2307/j.ctv1ghv45t>
- Crawford K, Paglen T (2019) Excavating AI: the politics of images in machine learning training sets. *Excavating AI*
- Dahlbäck N, Jönsson A, Ahrenberg L (1993) Wizard of Oz studies—why and how. *Knowl-Based Syst* 6(4):258–266. [https://doi.org/10.1016/0950-7051\(93\)90017-N](https://doi.org/10.1016/0950-7051(93)90017-N)
- de Reuver M, van Wynsberghe A, Janssen M, van de Poel I (2020) Digital platforms and responsible innovation: expanding value sensitive design to overcome ontological uncertainty. *Ethics Inf Technol* 22(3):257–267. <https://doi.org/10.1007/s10676-020-09537-z>
- Devendorf L, Ryokai K (2015) Being the machine: reconfiguring agency and control in hybrid fabrication. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp 2477–2486
- DiSalvo C (2015) *Adversarial design*. MIT Press, Cambridge
- Dove G, Halskov K, Forlizzi J, Zimmerman J (2017) UX design innovation: challenges for working with machine learning as a design material. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, CHI '17, pp 278–288, <https://doi.org/10.1145/3025453.3025739>
- EdgeImpulse (2019) Edge Impulse. <https://www.edgeimpulse.com/>
- Ehsan U, Riedl MO (2020) Human-centered explainable AI: towards a reflective sociotechnical approach. In: Stephanidis C, Kurosu M, Degen H, Reinerman-Jones L (eds) *HCI International 2020—Late Breaking Papers: Multimodality and Intelligence*, Springer International Publishing, Cham, Lecture Notes in Computer Science, pp 449–466, https://doi.org/10.1007/978-3-030-60117-1_33
- Elsden C, Chatting D, Durrant AC, Garbett A, Nissen B, Vines J, Kirk DS (2017) On speculative enactments. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, pp 5386–5399
- Elwes J (2019) Zizi. <https://www.jakeelwes.com/project-zizi-2019.html>
- Fiebrink R (2019) Machine learning education for artists, musicians, and other creative practitioners. *ACM Trans Comput Educ* 19(4):31:1–31:32. <https://doi.org/10.1145/3294008>
- Fiebrink R (2022) Machine learning for musicians and artists. <https://www.kadenze.com/courses/machine-learning-for-musicians-and-artists-v/info>
- Findeli A (2001) Rethinking design education for the 21st century: theoretical, methodological, and ethical discussion. *Des Issues* 17(1):5–17
- Fish B, Stark L (2021) Reflexive design for fairness and other human values in formal models. In: Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, Association for Computing Machinery, New York, NY, USA, pp 89–99
- Floridi L (2020) AI and its new winter: from myths to realities. *Philos Technol* 33(1):1–3. <https://doi.org/10.1007/s13347-020-00396-6>
- Forlano L, Mathew A (2014) From design fiction to design friction: speculative and participatory design of values-embedded urban technology. *J Urban Technol* 21(4):7–24
- Frauenberger C (2020) Entanglement HCI the next wave? *ACM Trans Comput-Hum Interact* 27(1):1–27. <https://doi.org/10.1145/3364998>
- Gebru T, Morgenstern J, Vecchione B, Vaughan JW, Wallach H, Daumé III H, Crawford K (2018) Datasheets for datasets. *arXiv preprint arXiv:1803.09010*
- Ghajargar M, Bardzell J (2023) Making AI understandable by making it tangible: exploring the design space with ten concept cards. In: Proceedings of the 34th Australian Conference on Human-Computer Interaction, Association for Computing Machinery,

- New York, NY, USA, OzCHI '22, pp 74–80, <https://doi.org/10.1145/3572921.3572942>
- Ghajargar M, Bardzell J, Renner AS, Krogh PG, Höök K, Cuartielles D, Boer L, Wiberg M (2021) From "Explainable AI" to "Graspable AI". In: Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction, Association for Computing Machinery, New York, NY, USA, TEI '21, pp 1–4, <https://doi.org/10.1145/3430524.3442704>
- Ghajargar M, Bardzell J, Smith-Renner AM, Höök K, Krogh PG (2022) Graspable AI: physical forms as explanation modality for explainable AI. In: Sixteenth International Conference on Tangible, Embedded, and Embodied Interaction, Association for Computing Machinery, New York, NY, USA, TEI '22, pp 1–4, <https://doi.org/10.1145/3490149.3503666>
- Giaccardi E (2019) Histories and futures of research through design: from prototypes to connected things. *Int J Des* 13(3):139–155
- Giaccardi E, Redström J (2020) Technology and more-than-human design. *Des Issues* 36(4):33–44. https://doi.org/10.1162/desi_a_00612
- Giaccardi E, Cila N, Speed C, Caldwell M (2016) Thing ethnography: doing design research with non-humans. In: Proceedings of the 2016 ACM Conference on Designing Interactive Systems, Association for Computing Machinery, New York, NY, USA, DIS '16, pp 377–387, <https://doi.org/10.1145/2901790.2901905>
- Hagendorff T (2020) The ethics of AI ethics: an evaluation of guidelines. *Mind Mach* 30(1):99–120. <https://doi.org/10.1007/s11023-020-09517-8>
- Hemment D, Murray-Rust D, Belle V, Aylett R, Vidmar M, Broz F (2022a) Experiential AI: enhancing explainability in artificial intelligence through artistic practice
- Hemment D, Zeilinger M, Vidmar M, Elwes J, Warner H, Sarmiento D, Hill R (2022b) Towards a heuristic model for experiential AI: analysing the Zizi show in the new real. In: DRS Biennial Conference Series, <https://dl.designresearchsociety.org/drs-conference-papers/drs2022/researchpapers/311>
- Hildebrandt M (2020) Smart technologies. *Internet Policy Rev* 9(4):1–16
- Hodder I (2016) Studies in human-thing entanglement. Online
- Holton R, Boyd R (2021) 'Where are the people? What are they doing? Why are they doing it?' (Mindell) Situating artificial intelligence within a socio-technical framework. *J Sociol* 57(2):179–195. <https://doi.org/10.1177/1440783319873046>
- HuggingFace (2016) Hugging face—The AI community building the future. <https://huggingface.co/>
- Inkpen K, Chancellor S, De Choudhury M, Veale M, Baumer EPS (2019) Where is the human? Bridging the gap between AI and HCI. In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, CHI EA '19, pp 1–9, <https://doi.org/10.1145/3290607.3299002>
- Irani L, Silberman MS (2014) From critical design to critical infrastructure: lessons from turkopticon. *Interactions* 21(4):32–35. <https://doi.org/10.1145/2627392>
- Jonsson M, Tholander J (2022) Cracking the code: co-coding with AI in creative programming education. In: Creativity and Cognition, Association for Computing Machinery, New York, NY, USA, C & C '22, pp 5–14, <https://doi.org/10.1145/3527927.3532801>
- Jung JY, Qiu S, Bozzon A, Gadiraju U (2022) Great chain of agents: the role of metaphorical representation of agents in conversational crowdsourcing. In: CHI Conference on Human Factors in Computing Systems, ACM, New Orleans LA USA, pp 1–22, <https://doi.org/10.1145/3491102.3517653>
- Khadpe P, Krishna R, Fei-Fei L, Jeffrey Hancock T, Bernstein Michael S (2020) Conceptual metaphors impact perceptions of human-AI collaboration. In: Proceedings of the ACM on Human-Computer Interaction. <https://doi.org/10.1145/3415234>
- Kirman B, Bates O, Lord C, Alter H (2022) Thinking outside the bag: worker-led speculation and the future of gig economy delivery platforms. In: DRS Biennial Conference Series, <https://dl.designresearchsociety.org/drs-conference-papers/drs2022/researchpapers/107>
- Kliman-Silver C, Siy O, Awadalla K, Lentz A, Convertino G, Churchill E (2020) Adapting user experience research methods for AI-driven experiences. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp 1–8
- Kovaleva O, Romanov A, Rogers A, Rumshisky A (2019) Revealing the dark secrets of BERT. <https://doi.org/10.48550/arXiv.1908.08593>, arXiv:1908.08593
- Lin CH (2014) Rethinking design education—a study for the importance of general education in design. *Int J Lib Arts Soc Sci* 2(9):88–97
- Lindgren S, Holmström J (2020) A social science perspective on artificial intelligence: building blocks for a research agenda. *J Dig Soc Res* 2(3):1–15. <https://doi.org/10.33621/jdsr.v2i3.65>
- Littman ML (2021) Gathering strength, gathering storms: the one hundred year study on artificial intelligence (AI100). Tech. rep., Stanford University, <https://ai100.stanford.edu/2021-report/gathering-strength-gathering-storms-one-hundred-year-study-artificial-intelligence>
- Liu V, Chilton LB (2022) Design guidelines for prompt engineering text-to-image generative models. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, CHI '22, pp 1–23, <https://doi.org/10.1145/3491102.3501825>
- Llano MT, d'Inverno M, Yee-King M, McCormack J, Ilisar A, Pease A, Colton S (2022) Explainable computational creativity. <https://doi.org/10.48550/arXiv.2205.05682>, arXiv:2205.05682
- Lloyd P (2019) You make it and you try it out: seeds of design discipline futures. *Des Stud* 65:167–181. <https://doi.org/10.1016/j.destud.2019.10.008>
- Lockton D, Singh D, Sabnis S, Chou M, Foley S, Pantoja A (2019) New metaphors: a workshop method for generating ideas and reframing problems in design and beyond. In: Proceedings of the 2019 on Creativity and Cognition, Association for Computing Machinery, New York, NY, USA, C & C '19, p 319–332, <https://doi.org/10.1145/3325480.3326570>
- Lupetti ML, Smit I, Nicenboim I, Huisman G, Villaronga EF, Bedo V, Jongeling M (2020) Subversive citizen manual for the more-than-human city
- Lyons MJ (2020) Excavating "Excavating AI": the elephant in the gallery. <https://doi.org/10.5281/zenodo.4037538>, arXiv:2009.01215
- Malsattar N, Kihara T, Giaccardi E (2019) Designing and prototyping from the perspective of AI in the wild. In: Proceedings of the 2019 on Designing Interactive Systems Conference, Association for Computing Machinery, New York, NY, USA, DIS '19, pp 1083–1088, <https://doi.org/10.1145/3322276.3322351>
- Marenko B, van Allen P (2016) Animistic design: how to reimagine digital interaction between the human and the nonhuman. *Dig Creat* 27(1):52–70. <https://doi.org/10.1080/14626268.2016.1145127>
- McCardle JR (2002) The challenge of integrating AI & smart technology in design education. *Int J Technol Des Educ* 12(1):59–76. <https://doi.org/10.1023/A:1013089404168>
- McCormack J, Hutchings P, Gifford T, Yee-King M, Llano MT, Dinverno M (2020) Design considerations for real-time collaboration with creative artificial intelligence. *Organ Sound* 25(1):41–52. <https://doi.org/10.1017/S1355771819000451>

- Mital P (2016) Creative applications of deep learning. <https://www.kadenze.com/courses/creative-applications-of-deep-learning-with-tensorflow-iv/info>
- Mital P (2021) Cultural appropriation with machine learning. <http://pkmital.com/home/2021/09/01/ucla-course-on-cultural-appropriation-with-machine-learning/>
- Mitchell M, Wu S, Zaldivar A, Barnes P, Vasserman L, Hutchinson B, Spitzer E, Raji ID, Gebru T (2019) Model cards for model reporting. In: Proceedings of the Conference on Fairness, Accountability, and Transparency, Association for Computing Machinery, New York, NY, USA, FAT* '19, pp 220–229, <https://doi.org/10.1145/3287560.3287596>
- Mohammad SM (2021) Ethics sheets for AI tasks. [arXiv:2107.01183](https://arxiv.org/abs/2107.01183) [cs] 2107.01183
- Murray-Rust D, Tsiakas K (2022) EthicAmanuensis: supporting machine learning practitioners making and recording ethical decisions. In: International Conference on Tool with Artificial Intelligence
- Murray-Rust D, Gorkovenko K, Burnett D, Richards D (2019) Entangled ethnography: towards a collective future understanding. In: Proceedings of the Halfway to the Future Symposium 2019, Association for Computing Machinery, New York, NY, USA, HTTF 2019, pp 1–10, <https://doi.org/10.1145/3363384.3363405>
- Murray-Rust D, Elsdon C, Nissen B, Tallyn E, Pschetz L, Speed C (2022a) Blockchain and beyond: understanding blockchains through prototypes and public engagement. Transactions on Computer-Human Interaction [arXiv:2112.11891](https://arxiv.org/abs/2112.11891)
- Murray-Rust D, Nicenboim I, Lockton D (2022b) Metaphors for designers working with AI. In: DRS Biennial Conference Series, <https://doi.org/10.21606/drs.2022.667>
- Natale S, Ballatore A (2020) Imagining the thinking machine: technological myths and the rise of artificial intelligence. *Convergence* 26(1):3–18. <https://doi.org/10.1177/1354856517715164>
- Nicenboim I, Giaccardi E, Søndergaard MLJ, Reddy AV, Strengers Y, Pierce J, Redström J (2020) More-than-human design and AI: in conversation with agents. In: Companion Publication of the 2020 ACM Designing Interactive Systems Conference, ACM, Eindhoven Netherlands, pp 397–400, <https://doi.org/10.1145/3393914.3395912>
- Nicenboim I, Giaccardi E, Redström J (2022) From explanations to shared understandings of AI. DRS Biennial Conference Series
- Nicolini D (2009) Zooming in and zooming out: a package of method and theory to study work practices. In: Worline M (ed) *Organizational ethnography: studying the complexities of everyday life*. SAGE Publications Ltd., pp 120–138
- Noor AK (2017) AI and the future of the machine design. *Mech Eng* 139(10):38–43. <https://doi.org/10.1115/1.2017-Oct-2>
- OpenCV (1999) OpenCV. <https://opencv.org/>
- PAIR (2020) People + AI Guidebook. <https://pair.withgoogle.com/guidebook>
- Pierce J (2021) Eccentric sensing devices: using conceptual design notes to understand design opportunities, limitations, and concerns connected to digital sensing. *Creativity and cognition*, vol 37. Association for Computing Machinery, New York, pp 1–1
- Pschetz L, Bastian M (2018) Temporal design: rethinking time in design. *Des Stud* 56:169–184. <https://doi.org/10.1016/j.destud.2017.10.007>
- Pschetz L, Pothong K, Speed C (2019) Autonomous distributed energy systems: problematising the invisible through design, drama and deliberation. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, ACM, Glasgow Scotland UK, pp 1–14, <https://doi.org/10.1145/3290605.3300617>
- Rapp A (2021) Wearable technologies as extensions: a postphenomenological framework and its design implications. *Hum-Comput Interact*. <https://doi.org/10.1080/07370024.2021.1927039>
- Reddy A, Nicenboim I, Pierce J, Giaccardi E (2020) Encountering ethics through design: a workshop with nonhuman participants. *AI & SOC*. <https://doi.org/10.1007/s00146-020-01088-7>
- Reddy A, Kocaballi AB, Nicenboim I, Søndergaard MLJ, Lupetti ML, Key C, Speed C, Lockton D, Giaccardi E, Grommé F, Robbins H, Primlani N, Yurman P, Sumartojo S, Phan T, Bedö V, Strengers Y (2021) Making everyday things talk: speculative conversations into the future of voice interfaces at home. In: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, CHI EA '21, pp 1–16, <https://doi.org/10.1145/3411763.3450390>
- Redström J, Wiltse H (2018) *Changing things : the future of objects in a digital world*. Bloomsbury Academic
- Runway (2020) Runway—Online Video Editor | Everything you need to edit video, fast. <https://runwayml.com>
- Samoli S, López-Cobo M, Gómez E, De Prato G, Martínez-Plumed F, Delipetrev B (2020) AI watch: defining artificial intelligence: towards an operational definition and taxonomy of artificial intelligence. Tech. rep., Joint Research Centre, European Commission, <https://data.europa.eu/doi/10.2760/382730>
- Sanches P, Howell N, Tsaknaki V, Jenkins T, Helms K (2022) Diffraction-in-action: designerly explorations of agential realism through lived data. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, CHI '22, pp 1–18, <https://doi.org/10.1145/3491102.3502029>
- Sanders EBN, Stappers PJ (2014) Probes, toolkits and prototypes: three approaches to making in codesigning. *CoDesign* 10(1):5–14. <https://doi.org/10.1080/15710882.2014.888183>
- Sartori L, Theodorou A (2022) A sociotechnical perspective for the future of AI: narratives, inequalities, and human control. *Ethics Inf Technol* 24(1):4. <https://doi.org/10.1007/s10676-022-09624-3>
- Scurto H, Caramiaux B, Bevilacqua F (2021) Prototyping machine learning through diffractive art practice. In: Designing Interactive Systems Conference 2021, Association for Computing Machinery, New York, NY, USA, DIS '21, pp 2013–2025, <https://doi.org/10.1145/3461778.3462163>
- Seymour W, Van Kleek M, Binns R, Murray-Rust D (2022) Respect as a lens for the design of AI systems. In: Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society, pp 641–652, <https://doi.org/10.1145/3514094.3534186>, [arXiv:2206.07555](https://arxiv.org/abs/2206.07555)
- Shen H, Deng WH, Chattopadhyay A, Wu ZS, Wang X, Zhu H (2021) Value cards: an educational toolkit for teaching social impacts of machine learning through deliberation. In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, ACM, Virtual Event Canada, pp 850–861, <https://doi.org/10.1145/3442188.3445971>
- Shotter J (2014) Agential realism, social constructionism, and our living relations to our surroundings: sensing similarities rather than seeing patterns. *Theory Psychol* 24(3):305–325. <https://doi.org/10.1177/0959354313514144>
- Siebert LC, Lupetti ML, Aizenberg E, Beckers N, Zgonnikov A, Veluwenkamp H, Abbink D, Giaccardi E, Houben GJ, Jonker CM, van den Hoven J, Forster D, Lagendijk RL (2022) Meaningful human control: actionable properties for AI system development. *AI and Ethics*. <https://doi.org/10.1007/s43681-022-00167-3>
- Simeone L, Mantelli R, Adamo A (2022) Pushing divergence and promoting convergence in a speculative design process: considerations on the role of AI as a co-creation partner. In: DRS Biennial Conference Series
- Sinders C, Ahmad S (2021) The labor behind the tools: using design thinking methods to examine content moderation software. *Interactions* 28(4):6–8. <https://doi.org/10.1145/3470492>

- Singh J, Cobbe J, Norval C (2019) Decision provenance: harnessing data flow for accountable systems. *IEEE Access* 7:6562–6574. <https://doi.org/10.1109/ACCESS.2018.2887201>
- Stappers PJ, Giaccardi E (2017) Research through design. In: *The Encyclopedia of Human-Computer Interaction*, The Interaction Design Foundation, <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/research-through-design>
- Stoffregen TA (2003) Affordances as properties of the animal-environment system. In: *how shall affordances be refined? Four perspectives*. Routledge
- Subramonyam H, Seifert C, Adar E (2021a) Protoai: model-informed prototyping for AI-powered interfaces. In: *26th International Conference on Intelligent User Interfaces*, pp 48–58
- Subramonyam H, Seifert C, Adar E (2021b) Towards a process model for co-creating AI experiences. *Des Interact Syst Conf* 2021:1529–1543. <https://doi.org/10.1145/3461778.3462012>
- Suchman L (2020) *Agencies in technology design: feminist reconfigurations*. Machine ethics and robot ethics. Routledge, London, pp 361–375
- TensorFlow (2015) TensorFlow. <https://www.tensorflow.org/>
- Theodorou A, Dignum V (2020) Towards ethical and socio-legal governance in AI. *Nat Mach Intell* 2(1):10–12. <https://doi.org/10.1038/s42256-019-0136-y>
- Tremblay PA, Roma G, Green O (2021) Enabling programmatic data mining as musicking: the fluid corpus manipulation toolkit. *Comput Music J* 45(2):9–23. https://doi.org/10.1162/comj_a_00600
- Troiano GM, Wood M, Sonbudak MF, Padte RC, Harteveld C (2021) "Are we now post-COVID?": exploring post-COVID futures through a gamified story completion method. In: *Designing Interactive Systems Conference 2021*, Association for Computing Machinery, New York, NY, USA, pp 48–63. <https://doi.org/10.1145/3461778.3462069>
- Vagg J (2022) Experiencing-with data: exploring posthuman creativity through rhizomatic empathy. *Qual Inq* 28(5):541–551. <https://doi.org/10.1177/10778004211069696>
- van der Burg V, Salah AA, Chandrasegaran S, Lloyd P (2022) Ceci n'est pas une chaise: emerging practices in designer-AI collaboration. *DRS Biennial Conference Series*
- Van der Helm A, Stappers PJ (2020) Prototyping with a team: acting machines support shared understanding. In: *DS 104: Proceedings of the 22nd International Conference on Engineering and Product Design Education (E & PDE 2020)*, VIA Design, VIA University in Herning, Denmark. 10th–11th September 2020
- van der Vlist B, van de Westelaken R, Bartneck C, Hu J, Ahn R, Barakova E, Delbressine F, Feijs L (2008) Teaching machine learning to design students. In: Pan Z, Zhang X, El Rhalibi A, Woo W, Li Y (eds) *Technologies for E-Learning and Digital Entertainment*, Springer, Berlin, Heidelberg, *Lecture Notes in Computer Science*, pp 206–217. https://doi.org/10.1007/978-3-540-69736-7_23
- Verbeek PP (2008) Cyborg intentionality: rethinking the phenomenology of human-technology relations. *Phenomenol Cogn Sci* 7(3):387–395. <https://doi.org/10.1007/s11097-008-9099-x>
- Verbeek PP (2020) Politicizing Postphenomenology. In: Miller G, Shew A (eds) *Reimagining philosophy and technology, reinventing ihde, philosophy of engineering and technology*. Springer International Publishing, Cham, pp 141–155. https://doi.org/10.1007/978-3-030-35967-6_9
- Wang D, Churchill E, Maes P, Fan X, Shneiderman B, Shi Y, Wang Q (2020) From human–human collaboration to human–AI collaboration: Designing AI systems that can work together with people. In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp 1–6
- Wegener FE, Gonçalves MG, Dankfort Z (2019) Reflection-in-action when designing organizational processes: prototyping workshops for collective reflection-in-action. In: *Proceedings of the Design Society: International Conference on Engineering Design* 1(1):1255–1264. <https://doi.org/10.1017/dsi.2019.131>, <https://www.cambridge.org/core/journals/proceedings-of-the-international-conference-on-engineering-design/article/reflectioninaction-when-designing-organizational-processes-prototyping-works-hops-for-collective-reflectioninaction/E7E5E6EB5924415E90A585472389FDA8>
- Westerlund B, Wetter-Edman K (2017) Dealing with wicked problems, in messy contexts, through prototyping. *Des J* 20(sup1):S886–S899. <https://doi.org/10.1080/14606925.2017.1353034>
- Wiltse H (ed) (2020) *Relating to things: design*. In: *Technology and the artificial*. Bloomsbury Publishing Plc. <https://doi.org/10.5040/9781350124288>
- Wong JS (2018) Design and fiction: imagining civic AI. *Interactions* 25(6):42–45. <https://doi.org/10.1145/3274568>
- Wong RY, Van Wyk E, Pierce J (2017) Real-fictional entanglements: using science fiction and design fiction to interrogate sensing technologies. In: *Proceedings of the 2017 Conference on Designing Interactive Systems*, pp 567–579
- Yang Q, Steinfeld A, Rosé C, Zimmerman J (2020) Re-examining whether, why, and how human-AI interaction is uniquely difficult to design. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp 1–13
- Yanow D, Tsoukas H (2009) What is reflection-in-action? A phenomenological account. *J Manag Stud* 46(8):1339–1364. <https://doi.org/10.1111/j.1467-6486.2009.00859.x>
- Yildirim N, Kass A, Tung T, Upton C, Costello D, Giusti R, Lacin S, Lovic S, O'Neill JM, Meehan RO (2022) How experienced designers of enterprise applications engage AI as a design material. In: *CHI Conference on Human Factors in Computing Systems*, pp 1–13
- Yildirim N, Oh C, Sayar D, Brand K, Challa S, Turri V, Crosby Walton N, Wong AE, Forlizzi J, McCann J, Zimmerman J (2023) Creating design resources to scaffold the ideation of AI concepts. In: *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, Association for Computing Machinery, New York, NY, USA, *DIS '23*, pp 2326–2346. <https://doi.org/10.1145/3563657.3596058>
- Yun G, Cho K, Jeong Y, Nam TJ (2022) Ideasquares: utilizing generative text as a source of design inspiration. *DRS Biennial Conference Series*
- Yurrita M, Agate B, Murray-Rust D, Bozzon A (2022) Towards a multi-stakeholder value-based assessment framework for algorithmic systems. In: *FACCT2022*

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.