Machine Learning Explanations as Boundary Objects: How AI Researchers Explain and Non-Experts Perceive Machine Learning

Amid Ayobi[a], Katarzyna Stawarz[b], Dmitri Katz[c], Paul Marshall[a], Taku Yamagata[a], Raúl Santos-Rodríguez[a], Peter Flach[a], and Aisling Ann O'Kane[a]

[a] *University of Bristol, Bristol, England*
[b] *Cardiff University, Cardiff, Wales*
[c] *The Open University, Milton Keynes, England*

Abstract

Understanding artificial intelligence (AI) and machine learning (ML) approaches is becoming increasingly important for people with a wide range of professional backgrounds. However, it is unclear how ML concepts can be effectively explained as part of human-centred and multidisciplinary design processes. We provide a qualitative account of how AI researchers explained and non-experts perceived ML concepts as part of a co-design project that aimed to inform the design of ML applications for diabetes self-care. We identify benefits and challenges of explaining ML concepts with analogical narratives, information visualisations, and publicly available videos. Co-design participants reported not only gaining an improved understanding of ML concepts but also highlighted challenges of understanding ML explanations, including misalignments between scientific models and their lived self-care experiences and individual information needs. We frame our findings through the lens of Stars and Griesemer's concept of boundary objects to discuss how the presentation of user-centred ML explanations could strike a balance between being plastic and robust enough to support design objectives and people's individual information needs.

Keywords 1

Explainable AI, AI literacy, Explanation, Diabetes, Boundary Objects

1. Introduction and Related Work

Understanding artificial intelligence (AI) approaches is becoming increasingly important for industry practitioners with a wide range of professional backgrounds and academic researchers working in interdisciplinary fields, such as human-computer interaction (HCI). While HCI and AI research have often been characterised as having quite distinct views of the relationship between humans and technology [30], more recent work has sought to integrate the approaches drawing not only on human-centred but also participatory HCI methodologies to understanding both how AI technology is being developed and how human-AI interactions could be designed. *"What I do know is that the future is not AI; it can only be an AI enabled through HCI,"* writes Harper [12], reflecting on the important role HCI could play in the new age of AI. In particular, the HCI community has looked at practices of researchers, data scientists, user experiences

Joint Proceedings of the ACM IUI 2021 Workshops, April 13-17, 2021, College Station, USA
EMAIL: amid.ayobi@bristol.ac.uk (A. 1); stawarzk@cardiff.ac.uk (A. 2); dmitrikatz23@gmail.com (A. 3); p.marshall@bristol.ac.uk (A. 4); taku.yamagata@bristol.ac.uk (A. 5), enrsr@bristol.ac.uk (A. 6); peter.flach@bristol.ac.uk (A. 7); a.okane@bristol.ac.uk (A. 8)
ORCID: 0000-0003-1104-0043 (A. 1); 0000-0001-9021-0615 (A. 2); 0000-0003-1345-7539 (A. 3); 0000-0003-2950-8310 (A. 4); 0000-0001-8624-7669 (A. 5); 0000-0001-9576-3905 (A. 6); 0000-0001-6857-5810 (A. 7); 0000-0001-8219-8126 (A. 8)

CEUR Workshop Proceedings (CEUR-WS.org)

designers, and end-users to bridge gaps between HCI and AI.

Pointing out that the manual work and human factors of ML research can be overlooked, Gillies et al. [10] and Clarke et al. [5] encourage researchers to draw on human-centred approaches to investigate the situated and collaborative facets of ML practices and the design of usable ML support tools. Taking up this call, Muller et al. [22] unpack how data scientists develop an intuitive sense of their datasets and how they create ground truth values as part of their data work. However, this perceived agency of working with data also has its limits. For example, based on a contextual inquiry, Kaur et al. [15] find that data scientists over-trust ML interpretability tools and face challenges to accurately describe output data visualisations.

As ML plays an increasingly important role in the design of products, not only data scientists but also designers engage with ML [17]. However, designing human-AI interactions entails major challenges [6, 7, 11, 31, 32]. For example, design professionals report difficulties in understanding ML capabilities, and recommend adopting data science jargon, including the use of quantitative evaluation methods, to be able to contribute to a data-centric work culture [31]. Envisioning a variety of feasible AI experiences and rapidly prototyping realistic human-AI interactions are further challenges that designers are faced with, considering time extensive ML training workflows and a lack of data to design with [6, 32, 33]. Furthermore, designers can find it difficult to productively collaborate with AI engineers because of a lack of a shared language and methodologies that help align human-cantered design and machine learning work streams [11].

Moving on from how data scientists and designers work with AI concepts and tools, prior work has drawn on participatory approaches to investigate end-users' perceptions and the ethical implications of AI systems [9, 21, 23, 28]. In particular, Loi et al. [17, 18] have highlighted that participatory design approaches are suitable to address AI challenges and inform AI futures: participatory design has been shown to be a powerful methodology to explore the design space of desirable technologies and foster mutual learning between multidisciplinary actors [2, 24, 26, 29]. For example, Katan et al. [13] have demonstrated the utility of interactive machine learning to support people with disabilities in creating and customising gesturally controlled musical interfaces through a series of participatory design workshops. Although participants faced challenges in understanding the training process to build instruments, they managed to appropriate pre-trained instruments according to their capabilities.

2. Method

The objective of this study was to investigate how ML explanations were presented and perceived as part of a co-design project that aimed to co-design ML-based decision support concepts and co-create suitable machine learning approaches. The project involved HCI researchers, AI researchers, and industry practitioners, as well as fifteen participants with T1 diabetes. This paper focuses on one workshop that specifically mediated ML concepts to workshop participants. We did not aim to evaluate the effectiveness or efficiency of the ML explanations. Instead, we investigated the following research questions:

- How did AI researchers explain ML concepts to co-design workshop participants?
- How did co-design workshop participants perceive the presented ML explanations?
- What are the transferable implications for designing user-centred ML explanations?

The first author conducted 18 interviews via phone and video conference systems. Interviews involved eight people with T1D who participated in the co-design project (referred to as P1, P2, etc.), three HCI researchers (e.g. HCI1), and three AI researchers (e.g. AI1). To support recollection before the interviews, a slide deck was shared with participants including ML explanations used throughout the workshop. Interview topics covered prior experiences with AI/ML and perceptions of ML explanations. Interview questions were adjusted for each group of interviewees and lasted approximately 30 minutes. The audio recordings were transcribed verbatim. This interview study received an ethical approval from the Faculty Ethics Committee.

Data collection and analysis was conducted in a staggered way according to project roles. A qualitative data analysis software was used by the first author to thematically code data [3]. As some participants were authors, each interviewee was sent the representative quotes for the codes and explicitly agreed to their use before group analysis was conducted. The data corpus was iteratively analysed in an inductive fashion drawing on open coding by all the authors [3].

3. Findings

We first report on how AI researchers explained ML concepts to participants as part of a co-design workshop using different types of explanations, including analogical narratives, data visualisations, and publicly available videos. We then describe how workshop participants, including HCI researchers and people with diabetes, perceived the presented ML explanations and what benefits and challenges they experienced.

3.1. ML Explanations

Since the objective of the co-design project involved the design of ML based applications for diabetes self-management, AI researchers used different methods to explain ML approaches to workshop participants, including data visualisations, analogies and videos of real-world AI applications.

3.1.1. Data Visualisation: Anomaly Detection

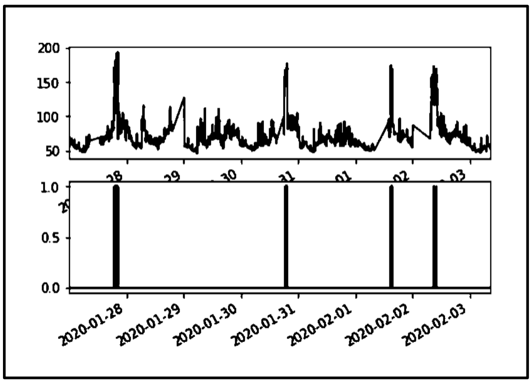

Figure 1: line graphs to explain anomaly detection

The concept of anomaly detection was explained with the help of two line graphs (see Figure 1). The first line graph showed continuous blood glucose measures over time in milligrams per decilitre. Representing a binary machine interpretation, the second line graph highlighted four anomalies in the continuous blood glucose data of the first line graph. Participants reported being used to reflect on line graphs when using different health and wellbeing applications [14]. However, they wished to hear narratives that described the real-word context and experiences of the person who collected the data to be able to relate and make sense of the anomaly explanation. For example, P8 made it clear that it is important not only to understand the contributing factors of anomalies but also how anomalies could be managed:

> "*What you're not really seeing is why those anomalies are happening. [...] if we're talking about diabetes, I think the 'why' is just as important in order to understand how to tackle those anomalies.*" (P8)

Moreover, participants highlighted that binary representations of anomalies (see Figure 1, second line graph) may be useful to explain the concept of anomaly detection, however, potentially not suitable to support sense-making and decision-making in everyday life. They felt more comfortable with data visualisations that supported their agency in identifying and dismissing anomalies based on their lived experience. For example, high blood glucose values in daily life were not necessarily anomalous if participants were able to make educated guesses about contributing contextual factors and manage these situations.

3.1.2. Analogy: Reinforcement Learning

Another ML concept that was explained as part of the co-design workshops was reinforcement learning. AI researchers mediated the concept of reinforcement learning with the help of the analogy of training a dog.

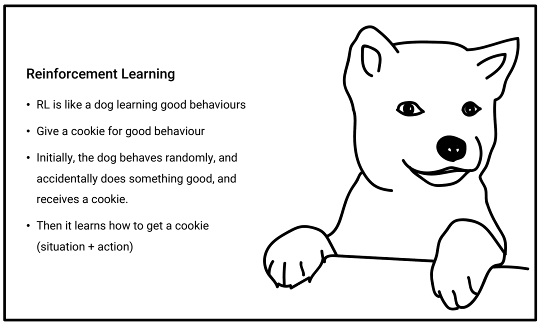

Figure 2: analogy to explain reinforcement learning

"*At first, it was a bit like, 'What!?' and then, when it was explained, it was like, 'Oh, yes, that makes sense,'*" P5 remembered, indicating that understanding this analogy requires translating the act of training a dog to the act of training a software agent that aims to maximise reward in a given environment. Participants reused the analogy of training a dog in different contexts, such as P8 who wished to be able to use a semi-automated self-tracking approach [4] that empowers people to manually stop false machine interpretations:

> *"So, you could use the dog example again, where it might be learning something which necessarily isn't correct, if that makes sense, like it might find a pattern which you don't want it to learn. So, I think... I don't think it's a question of like manually versus automatic. I think they need to work together in some shape or form. [...] there needs to be some sort of manual input to tell the machine learning aspect, 'Please don't learn this.'"* (P8)

Participants also perceived limitations of using the analogy of training a dog with cookies. For example, P3's account refers to the challenges of transferring anticipated emotions, such as the desire to learn, to machines and the challenges of translating the analogy to the design space of digital applications:

> "*If a machine has desire or it's how you explain the one for a cookie. I think that's the bit I find it hard to get my head round with a machine [...] So, I don't know how you reward an app like a machine*" (P3).

Furthermore, P10 politely critiqued the use of the term 'cookie' in the context of diabetes management, considering that cookies can be associated with dietary challenges people with diabetes can experience:

> *"I've got dogs and I give them treats, little dog treats. I think the use of the word cookie I found amusing shall we say. Because cookies are not a reward for us diabetics. In fact, that's a challenge."* (P10)

3.1.3. Video: Agent Behaviour

In addition, researchers used a seminal video [8], that is widely cited in the machine learning community, to demonstrate how agents learn to play the game of hide-and-seek. The video showed how agents developed strategies and counterstrategies over time, such as jumping on cubes and moving cubes to block doors. All participants described the video as a well-produced, powerful and memorable exemplar that mediated machine learning driven multi-agent behaviour with advanced character design and an entertaining narrative:

> "*The way the video showed how they sort of developed and how they learned was really clear, and the characters are quite cute, so I think it was quite funny as well, at the same time. Again, that was a great example to show how machine learning can work.*" (P5).

However, similar to the analogy of training a dog, participants found it challenging to transfer the hide-and-seek game to their diabetes self-management practices, highlighting that machine learning explanations need not only be abstracted but also transferred to a personally meaningful and research-specific context:

> "*I'm not sure how to transfer that to a diabetic situation in a way, that particular format. I mean there must be one, I haven't really thought that one through. [...] what have you got to have? You've got to have something whereby you're correlating eating or carb intake, exercise and taking insulin. So, those three factors, I think.*" (P6)

3.2. Understanding of ML Explanations

HCI researchers and workshop participants reported gaining an improved understanding of the presented ML approaches. Participants explained that even though they might not fully understand the "inner workings" (P3) of ML approaches, it was important to gain some knowledge of ML concepts to develop trust in the design process and potential ML implementations, though some noted the importance of it being presented in understandable terms:

> "*I don't think you just blindly follow stuff, particularly when designs are being made in the background […] it's better to put it into terms that we could understand, which is quite difficult when it can be so complex, but I do think it's quite important to give us some understanding of how and what's going on in the background*." (P5)

HCI researchers and participants reported that learning about ML approaches as part of the workshops changed their prior understanding of the benefits and limitations of ML based technologies. *"Before, it was kind of like, you know, computers being able to think for themselves or like have a sentience,"* P8 explained, exemplifying that some participants' prior understanding of AI was based on science fiction narratives that typically portray AI technologies with potentially dangerous autonomous and emotional capacities. Reflecting on their co-design workshop experiences, participants demonstrated differing degrees of ML literacy in creative ways. For example, participants used existing digital consumer services as examples to explain ML functionality, such as recommendations:

> "*I think the term 'artificial intelligence' is a bit more specific than that, I think. It's more to do with machine learning, […] So it's things like, you know, how Netflix decides what you watch, kind of thing, or how you choose your recommendation. I think it's algorithms, really."* (P3)

Participants described AI research and AI concepts, such as ML, as data driven algorithms that are written by humans and run on computers. *"AI is computers that learn, that once you set certain criteria up or whatever, they can gain knowledge themselves without being told to gain knowledge, yeah. I think that is the simplest form,"* explained P10, referencing the learning capabilities of AI technologies. Participants with diabetes also reflected on potential limitations of ML approaches, including differences between manual and automatic data collection, roles of data quality and potential limitations of predictive functionalities:

> *"If it's showing information based on weeks and weeks of data-gathering and it's basically giving you your average day, I mean, I suppose that could be useful. But then, if you suddenly change your physical activity, or you're eating something at a time that you don't usually eat something, then I guess that could disrupt it."* (P8).

4. Discussion

Understanding AI approaches is becoming increasingly important for people with a wide range of professional backgrounds in industrial and academic settings. We have provided a qualitative account of how AI researchers explained ML concepts to HCI researchers and people with diabetes as part of a co-design project that aimed to inform the design of ML applications for diabetes self-care. Here we discuss our findings through the lens of Stars and Griesemer's concept of boundary objects to outline how the presentation of user-centred ML explanations could strike a balance between being plastic and robust enough to support design objectives and people's individual information needs as part of multidisciplinary projects.

4.1. Framing ML Explanations as Boundary Objects

Star and Griesemer's [25] concept of boundary objects has been used as a theoretical lens to understand how various actors with different backgrounds, roles, and interests successfully collaborate as part of multidisciplinary endeavours. Boundary objects are artefacts that facilitate communication and collaboration between multiple actors and are defined as:

> *"objects which are both plastic enough to adapt to local needs and the constraints of the several parties employing them, yet robust enough to maintain a common identity across sites"* (ibid, p. 393).

In their study of how amateurs, professionals, and administrators collaborate in a museum setting, Star and Griesemer distinguish between four types of boundary objects: (1) *repositories* provide a central location where objects, such as samples, are systematically stored and are available for people to be used; (2) *ideal type* is an object, such as a diagram, that provides an abstracted representation that can be adapted by others; (3) *coincident boundaries* are objects, such as tailored maps: they are defined by common (geographical) boundaries but can have different contents, purposes, and styles; (4) *standardised forms* are boundary objects that are used as formal methods of communication across different actors. While these four types of boundary objects can be used in different ways and can have different meanings for different actors from different social worlds, they typically support communication and facilitate collaborations. Although boundary objects aim to resolve conflicts, they are not neutral. The creation of boundary objects requires carefully managing power relationships to avoid forced use of predefined representations that can cause systematic exclusion, discrimination, and injustice.

In our case, AI researchers used different types of ML explanations to support HCI researchers and people with diabetes in co-designing possible ML systems. To foster a shared understanding of ML concepts, they used analogical narratives to explain reinforcement learning, data visualisations to explain anomaly detection, and publicly available videos to explain multi-agent behaviour. These explanations can be characterised as *ideal types*, based on Star and Griesemer's types of boundary objects. Framing these ML explanations as boundary objects poses the question what the theory of boundary objects and the key properties of boundary objects - robustness and plasticity - imply for the design of ML explanations.

4.2. Balancing Robustness and Plasticity

While the robustness of a ML explanation can be described with features, such as being algorithmically correct and transferable to different research settings, the plasticity of a ML explanation can be associated with features, such as being adaptable to people's lived experiences, reflective capacities, and information needs. Design techniques, such as personalisation and customisation are particularly suitable to support people's individual needs and experiences of agency, such as sense of identify and ownership [1].

A robust and plastic enough ML explanation support actors, such as a co-designer, product manager, and end-user, in making sense of and acting on a ML explanation.

In our study, we have observed that participants made sense of ML explanations based on their prior knowledge of AI narratives and technologies, reused ML explanations, such as the analogy of training a dog, as part of co-design activities, and co-created mockups that visualised possible ML-based functionalities, such as predicting blood glucose values.

An important contributing factor for adopting a ML explanation was familiarity: participants particularly valued the analogical narrative of training a dog, since it seemed to help bridge the unknown concept of reinforcement learning and the known practice of training a dog. Barriers to adopting and using a ML explanation seemed to be a lack of abstraction and associations with people's lived self-care experiences.

4.3. Considering Sociocultural Contexts and Ethical Implications

The sociocultural underpinning of boundary objects suggests that co-designing a plastic and robust enough ML explanation involves not only representing a specific ML concept correctly and evaluating whether the ML explanation was correctly understood, but also gaining a holistic and non-judgemental understanding of how the ML explanation was appropriated and experienced within a certain context. For example, our qualitive inquiry has revealed the importance of tailoring general ML explanations to specific cases, such as self-managing diabetes, to avoid misalignments between people's lived experience and scientific concepts of ML.

Conceptualising ML explanations as boundary objects means to acknowledge that abstraction and ambiguity can lead to divergent viewpoints, misinterpretations, and misunderstandings. Our findings suggest that gaining a good enough understanding of ML explanations can support participants in developing trust in design processes, data collection and analysis technologies, and overarching research objectives. However, what a good enough understanding is and whether a good enough understanding of ML explanations and functionalities is ethically responsible depends on contextual factors, such as the sensitivity of a research setting. While participants with diabetes sketched predictive functionalities during co-design activities, AI researchers highlighted fundamental differences between the desirability and feasibility of ML-driven systems considering fatal implications of false predictions and recommendations in the case of continuous blood glucose monitoring and management.

4.4. Applying User Experience Design Methods

Developing a plastic and robust enough ML explanation can require an iterative and multidisciplinary design process with a detailed understanding of ML approaches, user groups, and the intended purpose of a ML explanation.

Considering that design methods and tools to facilitate co-design are recognised methodological contributions [2, 16], we encourage researchers and practitioners to explore the design space of "learner-centered" [19] ML explanations specifically for human-centred technology projects. Such design-led inquiries could explore how scientific ML explanations could be intertwined with people's lived self-care experiences and their information needs as co-designers. These explanation instruments could represent AI/ML at a layer of abstraction above specific algorithms and communicate not just of what AI/ML can do, but also what it cannot.

Content could be presented in engaging ways, as demonstrated by the creative presentation of AI as a monster metaphor [7], the use of tangible cards in the context of data protection regulations [20], and "inspirational bits" [27] that expose dynamic properties of sensors to allow designers to understand and experience the properties of technology that might be used in research and design projects.

5. Conclusion

We have provided a qualitative account of how AI researchers explained and non-experts perceived ML concepts as part of a co-design project that aimed to inform the design of ML applications for diabetes self-care.

We have identified benefits and challenges of explaining ML concepts with analogical narratives, information visualisations, and publicly available videos. Co-design participants reported not only gaining an improved understanding of ML concepts but also gaining trust in the co-design process of ML based technologies, data collection and analysis technologies, and overarching research objectives. However, co-design participants also highlighted challenges of understanding ML explanations, including misalignments between scientific models of ML and their lived self-care experiences and prior knowledge of AI and ML approaches.

Based on this understanding, we have framed our findings through the lens of Stars and Griesemer's concept of boundary objects to discuss how the presentation of user-centred ML explanations could maintain a delicate balance between being plastic and robust enough to support design objectives and people's individual information needs as part of multidisciplinary projects.

6. Acknowledgements

This project was funded by an Innovate UK Digital Catalyst Award - Digital Health. RSR is partially funded by the UKRI Turing AI Fellowship EP/V024817/1. Many thanks to all study participants and reviewers.

7. References

[1] Ayobi, A. et al. 2020. Trackly: A Customisable and Pictorial Self-Tracking App to Support Agency in Multiple Sclerosis Self-Care. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)* (2020).

[2] Bødker, S. et al. 2000. Co-operative Design—perspectives on 20 years with 'the Scandinavian IT Design Model.' *proceedings of NordiCHI* (2000), 22–24.

[3] Braun, V. and Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*. 3, 2 (Jan. 2006), 77–101. DOI:https://doi.org/10.1191/1478088706qp063oa.

[4] Choe, E.K. et al. 2017. Semi-Automated Tracking: A Balanced Approach for Self-Monitoring Applications. *IEEE Pervasive Computing*. 16, 1 (Jan. 2017), 74–84. DOI:https://doi.org/10.1109/MPRV.2017.18.

[5] Clarke, M.F. et al. 2019. Better Supporting Workers in ML Workplaces. *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing* (Austin TX USA, Nov. 2019), 443–448.

[6] Dove, G. et al. 2017. UX Design Innovation: Challenges for Working with Machine Learning as a Design Material. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver Colorado USA, May 2017), 278–288.

[7] Dove, G. and Fayard, A.-L. 2020. Monsters, Metaphors, and Machine Learning. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu HI USA, Apr. 2020), 1–17.

[8] Emergent Tool Use from Multi-Agent Interaction: 2019. *https://openai.com/blog/emergent-tool-use/*. Accessed: 2020-09-17.

[9] Fu, Z. and Zhou, Y. 2020. Research on human–AI co-creation based on reflective design practice. *CCF Transactions on Pervasive Computing and Interaction*. 2, 1 (Mar. 2020), 33–41. DOI:https://doi.org/10.1007/s42486-020-00028-0.

[10] Gillies, M. et al. 2016. Human-Centred Machine Learning. *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '16* (San Jose, California, USA, 2016), 3558–3565.

[11] Girardin, F. and Lathia, N. 2017. When User Experience Designers Partner with Data Scientists. *AAAI Spring Symposia* (2017).

[12] Harper, R.H.R. 2019. The Role of HCI in the Age of AI. *International Journal of Human–Computer Interaction*. 35, 15 (Sep. 2019), 1331–1344. DOI:https://doi.org/10.1080/10447318.2019.1631527.

[13] Katan, S. et al. 2015. Using Interactive Machine Learning to Support Interface Development Through Workshops with Disabled People. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15* (Seoul, Republic of Korea, 2015), 251–254.

[14] Katz, D.S. et al. 2018. Data, Data Everywhere, and Still Too Hard to Link: Insights from User Interactions with Diabetes Apps. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2018), 503:1-503:12.

[15] Kaur, H. et al. 2020. Interpreting Interpretability: Understanding Data Scientists' Use of Interpretability Tools for Machine Learning. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu HI USA, Apr. 2020), 1–14.

[16] Kensing, F. and Blomberg, J. 1998. Participatory design: Issues and concerns. *Computer Supported Cooperative Work (CSCW)*. 7, 3–4 (1998), 167–185.

[17] Loi, D. et al. 2019. Co-designing AI Futures: Integrating AI Ethics, Social Computing, and Design. *Companion Publication of the 2019 on Designing Interactive Systems Conference 2019 Companion - DIS '19 Companion* (San Diego, CA, USA, 2019), 381–384.

[18] Loi, D. et al. 2018. PD manifesto for AI futures. *Proceedings of the 15th*

Participatory Design Conference on Short Papers, Situated Actions, Workshops and Tutorial - PDC '18 (Hasselt and Genk, Belgium, 2018), 1–4.

[19] Long, D. and Magerko, B. 2020. What is AI Literacy? Competencies and Design Considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu HI USA, Apr. 2020), 1–16.

[20] Luger, E. et al. 2015. Playing the Legal Card: Using Ideation Cards to Raise Data Protection Issues within the Design Process. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15* (Seoul, Republic of Korea, 2015), 457–466.

[21] Morrison, C. et al. 2017. Imagining Artificial Intelligence Applications with People with Visual Disabilities using Tactile Ideation. *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility* (Baltimore Maryland USA, Oct. 2017), 81–90.

[22] Muller, M. et al. 2019. How Data Science Workers Work with Data: Discovery, Capture, Curation, Design, Creation. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19* (Glasgow, Scotland Uk, 2019), 1–15.

[23] Muller, M. and Liao, Q.V. 2017. Exploring AI Ethics and Values through Participatory Design Fictions. *Human Computer Interaction Consortium*. (2017).

[24] Slattery, P. et al. 2020. Research co-design in health: a rapid overview of reviews. *Health Research Policy and Systems*. 18, 1 (Dec. 2020), 17. DOI:https://doi.org/10.1186/s12961-020-0528-9.

[25] Star, S.L. and Griesemer, J.R. 1989. Institutional ecology,translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social studies of science*. 19, 3 (1989), 387–420.

[26] Steen, M. 2013. Co-design as a process of joint inquiry and imagination. *Design Issues*. 29, 2 (2013), 16–28.

[27] Sundström, P. et al. 2011. Inspirational bits: towards a shared understanding of the digital material. *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11* (Vancouver, BC, Canada, 2011), 1561.

[28] Trewin, S. et al. 2019. Considerations for AI fairness for people with disabilities. *AI Matters*. 5, 3 (Dec. 2019), 40–63. DOI:https://doi.org/10.1145/3362077.3362086.

[29] Vartiainen, H. et al. 2020. Machine learning for middle-schoolers: Children as designers of machine-learning apps. *2020 IEEE Frontiers in Education Conference (FIE)* (Uppsala, Sweden, Oct. 2020), 1–9.

[30] Winograd, T. 2006. Shifting viewpoints: Artificial intelligence and human–computer interaction. *Artificial Intelligence*. 170, 18 (Dec. 2006), 1256–1258. DOI:https://doi.org/10.1016/j.artint.2006.10.011.

[31] Yang, Q. et al. 2018. Investigating How Experienced UX Designers Effectively Work with Machine Learning. *Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18* (Hong Kong, China, 2018), 585–596.

[32] Yang, Q. et al. 2020. Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu HI USA, Apr. 2020), 1–13.

[33] Yang, Q. et al. 2019. Sketching NLP: A Case Study of Exploring the Right Things To Design with Language Intelligence. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19* (Glasgow, Scotland Uk, 2019), 1–12.