How Experienced Designers of Enterprise Applications Engage AI as a Design Material

Nur Yildirim
Carnegie Mellon University
Pittsburgh, PA, USA
yildirim@cmu.edu

Alex Kass
Teresa Tung
Accenture Labs
San Francisco, CA, USA

Connor Upton
Donnacha Costello
Robert Giusti
Sinem Lacin
Sara Lovic
James O'Neill
Rudi O'Reilly Meehan
Eoin Ó Loideáin
Azzurra Pini
Medb Corcoran
Jeremiah Hayes
Diarmuid Cahalane
Gaurav Shivhare
Luigi Castoro
Giovanni Caruso
Accenture
Dublin, Ireland

Changhoon Oh
James McCann
Jodi Forlizzi
John Zimmerman
Carnegie Mellon University
Pittsburgh, PA, USA

ABSTRACT

HCI research has explored *AI as a design material*, suggesting that designers can envision AI's design opportunities to improve UX. Recent research claimed that enterprise applications offer an opportunity for AI innovation at the user experience level. We conducted design workshops to explore the practices of experienced designers who work on cross-functional AI teams in the enterprise. We discussed how designers successfully work with and struggle with AI. Our findings revealed that designers can innovate at the system and service levels. We also discovered that making a case for an AI feature's return on investment is a barrier for designers when they propose AI concepts and ideas. Our discussions produced novel insights on designers' role on AI teams, and the boundary objects they used for collaborating with data scientists. We discuss the implications of these findings as opportunities for future research aiming to empower designers in working with data and AI.

CCS CONCEPTS

• **Human-centered computing** → **Interaction design process and methods**.

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04.
https://doi.org/10.1145/3491102.3517491

KEYWORDS

User experience, design, artificial intelligence, machine learning

ACM Reference Format:
Nur Yildirim, Alex Kass, Teresa Tung, Connor Upton, Donnacha Costello, Robert Giusti, Sinem Lacin, Sara Lovic, James O'Neill, Rudi O'Reilly Meehan, Eoin Ó Loideáin, Azzurra Pini, Medb Corcoran, Jeremiah Hayes, Diarmuid Cahalane, Gaurav Shivhare, Luigi Castoro, Giovanni Caruso, Changhoon Oh, James McCann, Jodi Forlizzi, and John Zimmerman. 2022. How Experienced Designers of Enterprise Applications Engage AI as a Design Material. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3491102.3517491

1 INTRODUCTION

Artificial intelligence (AI) plays an increasingly important role in the user experience (UX) of products and services. It automates menial tasks, helps people find something they want, and even offers new insights on how events in the world will unfold. In response, design researchers have started investigating the concept of *"AI as a design material"*, exploring how designers have conversations with AI in situations that might be improved by its use. Research shows AI creates a number of challenges for design practitioners. For example, designers often struggle to understand AI capabilities [21, 93]. This causes designers to both fail to recognize low hanging fruit, situations where a little AI might help [94], and to frequently envision things that exceed AI's current capabilities [93].

To help address this situation, the design community has explored data and AI through design-led inquiry, providing first-person accounts of envisioning new things that leverage data and AI [4, 9, 10, 91]. Some researchers created methods and guidelines for design-oriented data exploration [32, 46, 77], and for refining the user experience of AI systems [2, 31, 65]. Researchers also studied designers who successfully worked on AI systems, noticing how they work differently from other designers [92]. They observed that when these experienced designers envisioned new AI innovations, they called upon internalized abstractions of AI capabilities, and they used existing examples of the capability to communicate their ideas to others. Experienced designers developed close collaborations with data scientists, who played a critical role in helping to envision new ideas and prototype selected innovations.

The importance of the interplay between design and data science motivated design researchers to speculate that new boundary objects, artifacts which better support communication between these two disciplines [79], offer an effective path for increasing AI's design innovation [12, 14, 91]. In addition, they speculated that adaptive user interfaces (AUI) offer a great space where UX designers are likely to recognize situations where AI can help [94], and that enterprise applications offer a particularly rich space to discover these UI level opportunities for adaptation [97].

Our team (HCI and AI researchers, UX and service designers, and data scientists) set out to explore design-led AI innovation in the enterprise. We sought to understand how designers on AI teams envision and identify AI opportunities. We hoped to assess the claim that AUIs present an opportune yet overlooked area for designers to innovate UX with AI. We held a series of design workshops to understand designers' roles on cross-functional AI teams, and how they successfully work with AI.

The workshops produced several interesting findings, including:

(1) Designers bring more impact when innovating at the system and service levels as opposed to searching for AI opportunities at the user interface level of a project.
(2) In contrast to prior work, designers did not struggle to recognize opportunities where AI would create value for users. Instead, justifying the business value and the return on investment cost was a bigger barrier to AI's design innovation.
(3) Designers' co-location and collaboration with data scientists, and the use of boundary objects to communicate between design and data science, offered a successful path for engaging with data and AI as design materials.

This study makes two contributions. First, we present a case study that details how designers on cross-functional enterprise AI teams engage with data and AI as a design material. We draw attention to collaborative practices and the value designers bring to AI teams. Second, we advance our community's understanding of how data and AI impacts design practice, specifically for envisionment. We raise several open questions as directions for future research.

2 RELATED WORK

Our research draws from HCI research investigating how designers work with data and AI as a design material; and work on design and data science collaboration.

2.1 Data and AI as Design Materials

HCI and design researchers have investigated the challenges of working with data and AI as design materials, and the role of design in mitigating potential harms of data-driven algorithmic systems. Several studies reported challenges in prototyping the user experience of AI systems [21, 93]. In response, researchers have developed practitioner-facing AI tools, methods, guidelines, and design patterns to aid designers in accounting for AI systems' UX breakdowns [2, 3, 31, 51, 61, 65], such as planning for AI inference errors [38] or setting user expectations [45].

Parallel to these efforts, researchers have explored AI's risks and its societal consequences in perpetuating existing inequities and biases. The HCI community studied the issues around fairness, accountability, transparency, and ethics as they relate to AI systems [6, 25, 49, 75, 78, 85]. For example, researchers studied users' perceptions related to fairness in AI systems [5, 48, 89], and practitioners' challenges and needs around designing responsible AI [37, 84]. There is a growing body of work for providing processes and tools to support practitioners in developing fairer systems [20, 56, 80]. Empowering designers in sketching and envisioning with data and AI remains relatively under-investigated [93].

One strand of research has explored envisioning with data and AI through design-led inquiry, such as Research through Design [10, 22, 35, 52, 63], Design Fiction [76, 88] or Speculative Design [4, 8, 67], to provide a first-person account. For example, [9] investigated sensor data from a connected baby bottle as a material for designing bottle feeding experiences for parents; [91] ideated with NLP capabilities to envision novel concepts for an intelligent writing assistance. Another strand investigated design practitioners' experience with design innovation for data and AI. This type of work appears infrequently in the literature, and it might be impacted by how frequently designers only join AI projects towards the end, in order to solve the problem of human-AI interaction.

Designers seem to rarely be involved in problem setting or ideation [21]. Researchers report that designers struggle to envision novel and interesting uses of AI due to challenges in understanding what AI can and cannot do, often leading them to conceive of ideas that cannot be built [21, 93]. An interview study revealed that designers who had built a large set of "designerly abstractions" of AI's capabilities and the value AI generated for users were successful in engaging AI to improve UX, often through personalization or adaptation [92]. Most of these experienced designers worked for large, AI-focused companies where they could form ongoing, close collaborations with data scientists to leverage their technical expertise as a proxy in understanding what AI can do. Less discussed is how and when design practitioners envision AI opportunities.

Building on these research strands, HCI researchers have proposed a few directions to support envisioning with data and AI. Some researchers focused on improving designers' literacy around data and AI by helping designers to learn how AI functions [34, 44] or by sensitizing them to designerly abstractions of AI, such as taxonomies of AI capabilities [40, 92]. However, there is no consensus on what constitutes a "good enough" understanding of AI [55, 68]. Some research focused on creating design-oriented data exploration workflows [32, 46, 77] or prototyping toolkits [26, 57, 83] for designers to gain a felt sense of working with data and AI. Others argued for design processes beyond user-centered design, such as processes that focus on multiple user groups and stakeholders [28, 42, 50, 53, 62, 86]; processes that focus on close collaboration between design and data science [30, 43, 81]; or processes that seek for potential value in existing datasets and AI systems [7, 90].

Table 1: Our team consisting of researchers (R) and practitioners (P) who had many different roles and experience.

ID	Sess.	Professional Role	Exp.	Org.
R1	All	Principal Director/Fellow	10+yrs	Enterprise Lab
R2	1-5	Chief Technologist	10+yrs	Enterprise Lab
R3	All	HCI Researcher/Designer	10+yrs	University
R4	All	HCI Researcher/Designer	10+yrs	University
R5	4-6	HCI Researcher/Designer	10+yrs	University
R6	All	HCI Researcher/Designer	5-7 yrs	University
R7	1-3	HCI Researcher/Designer	5-7 yrs	University
P1	1-3	UX Designer	7-9 yrs	Enterprise
P2	1-3	UX Designer	7-9 yrs	Enterprise
P3	4-6	Design Research Lead	10+yrs	Enterprise
P4	4-6	Service Design Lead	5-7 yrs	Enterprise

ID	Sess.	Professional Role	Exp.	Org.
P5	4-6	Service Design Lead	5-7 yrs	Enterprise
P6	4-6	Service/UX Designer	3-5 yrs	Enterprise
P7	4-6	Data Designer	3-5 yrs	Enterprise
P8	4-6	Data Designer	10+yrs	Enterprise
P9	4-6	Data Designer	10+yrs	Enterprise
P10	4-6	Design Lead	10+yrs	Enterprise
P11	4-6	Group Design Director	10+yrs	Enterprise
P12	6	AI R&D Managing Director	10+yrs	Enterprise
P13	6	AI Research Engineer	10+yrs	Enterprise
P14	6	AI Research Principal	7-9 yrs	Enterprise
P15	6	Data Architect	3-5 yrs	Enterprise

Recent work exploring AI's design innovation has surfaced that designers often focus on complex uses of AI when envisioning, where they could instead focus on well-known AI capabilities that are likely to improve UX [21, 93]. AUIs present such an opportunity where designers are well positioned to identify frequent and repetitive user behaviors while generating user scenarios and wireframing transactional flows [94]. This seems especially true for enterprise applications where time saved by adaptation and the cost of a worker's time can be more easily measured [97]. Our work builds on these insights, deepening the exploration around how experienced designers engage AI to envision design opportunities.

2.2 Collaboration and Boundary Objects

There has been a growing interest in the HCI community around the cross-disciplinary collaboration throughout the AI lifecycle [37, 59, 66, 69, 95]. Investigation of design and data science collaboration shows that practitioners face challenges due to a lack of shared workflow or common language [30, 43, 93]. This is characterized as a gap between the two practices: designers envision AI ideas that are beyond the limits of existing AI capabilities and cannot be built, and data scientists build AI things that users do not want [91]. Moreover, AI experts can be a scarce resource for UX teams [92].

Boundary objects, information or resources used by collaborative teams to foster shared understanding [47, 79], can scaffold cross-disciplinary collaboration among AI practitioners and stakeholders [12, 14, 41, 70, 91]. Relatively little work has explored the use of boundary objects between design and data science practitioners on industry AI teams. One study reported that abstractions of AI capabilities and data visualizations served as boundary objects to facilitate conversations between UX and AI expertise [92]. Some HCI researchers reflecting on their own design process proposed using wireframes with data annotations as boundary objects [91, 94]. While there is a common desire for supporting collaboration in cross-functional AI teams, the types of boundary objects that might support design and data science collaboration remain unknown. Our work advances these prior efforts by focusing on design practitioners working as part of cross-functional AI teams.

3 METHOD

We wanted to understand how experienced design practitioners who regularly work with AI in the enterprise envision AI-driven interactions, identify opportunities for AUIs, and play different roles in cross-functional AI teams. Our interdisciplinary research team (N = 22) included academic researchers and industry researchers working at a company that develops AI-powered enterprise software for many industrial clients. We leveraged the internal network of the company to add experienced UX and service design practitioners to the team who regularly designed human-AI interactions in products and services.

We conducted design workshops and held discussions to investigate design practices. We chose design workshops for several reasons: 1) observations of current practice were impractical for confidentiality reasons; 2) individual interviews could not capture their practice as the work was spread across teams and unfolded over time; 3) conducting design workshops allowed practitioners to collectively reflect on their work and articulate their current practices, leading to co-discoveries within groups [71].

We conducted three workshops with each of two groups, for a total of six workshops. Each workshop had 8-18 practitioners and researchers. We also brought in data scientists and AI engineers for one of the workshop sessions. Participants had worked in professional practice for more than 3 years. Table 1 provides a summary of our research teams' composition, relevant experience, and the participants' involvement in workshop sessions.

Each workshop session lasted between 1-2 hours (10 hours in total). Workshops were held over video conference, due to the ongoing COVID-19 pandemic. Each workshop session focused on a different stage in the design process, roughly to correspond to the early phase (Discover, Define), mid-phase (Define, Develop), and late phase (Develop, Deliver) [16]. We asked practitioners to complete a prework activity prior to workshop sessions, which scaffolded the workshop activities described below:

Workshop 1. Participants created customer journey maps of their design process to detail their current workflow, tools, and stakeholders as prework. Researchers shared a short video presentation of resources from literature on AUIs, such as examples and design patterns [39, 94, 97]. Participants reflected on their recent enterprise design projects where they improved UX with AI.

Researchers probed whether, when, and how they envisioned AI-driven interactions. To assess the claims around AUIs being an opportune yet overlooked space for designers to leverage AI, we discussed whether they identified AUI opportunities in their work.

Workshop 2. Participants shared projects showcasing an AI experience they designed, and artifacts for prototyping with data and AI, such as wireframes. Researchers shared a case study on designing AI-driven adaptations [94, 98]. We discussed practices for prototyping AI-driven experiences. We discussed scenario and wireframe generation with data and AI to understand how practitioners represented things like data dependency and user labelling of data in interfaces.

Workshop 3. Participants reflected on the tools, methods, and processes they used to capture, document, and transfer their AI-driven designs. Researchers introduced the concept of boundary objects for collaboration [14, 79]. Designers and data scientists collectively discussed process with a focus on the role of design practitioners in AI teams, and artifacts for cross-disciplinary communication and collaboration. Participants were asked to articulate challenges and pain points and best practices for designing human-AI interactions. Following this session, we conducted an additional meeting with participants for debrief.

Workshop sessions were recorded and transcribed, artifacts generated during the sessions or shared prior or afterward were documented. We analyzed the transcripts using thematic analysis [11]. The first author conducted an inductive analysis, first open coding and then discussing themes with the research team. We then iteratively reviewed and refined the codes for synthesis. This produced 11 clusters, from which we constructed three main themes.

4 FINDINGS

We initially focused on how designers who work on enterprise applications recognize opportunities where an AUI might add value. However, we quickly realized AUIs played a minor role in their work. Instead, how the designers worked with and recognized opportunities to use data and AI became our central concern. Findings coalesced into three overlapping categories: AI as a design material, co-creating value and barriers to AI innovation, and collaboration with data scientists. We provide details on the designers' context and then describe the three categories.

Designers worked on cross-disciplinary AI teams, regularly working closely with data scientists, software developers, business managers, and other domain experts. They worked on internal and client-facing projects with development times ranging from 1-3 years. Designers followed a four-phase, double diamond design process [16]. In addition, they ran Agile Scrum design sprints within the phases [13]. They provided design support for internal projects, including the design of new enterprise applications, redesign of legacy applications, and design of new platforms that could be used across their company's clients. For example, they had recently designed the interface for an intelligent forecasting system. This platform would be used by many clients working in retail. It allowed business analysts to explore why predicted product sales might or might not match recent sales. They also worked directly for clients, working on first-of-a-kind projects that integrated AI in novel ways. For example, they had recently designed a human-in-the-loop logistics platform that performed dynamic route optimization by enabling workers to use their context knowledge and common sense to create detailed delivery plans [1].

4.1 AI as a Design Material

In response to questions about how they envision AI-driven experiences, designers shared several examples where they recognized opportunities for AI to improve a situation. Prior literature suggests that designers would discuss interface level product features and interactions, such as AUIs [92]. Interestingly, participants shared that UI-level opportunities only partly cover how they innovate with AI. In explaining how they worked, designers sketched a diagram to help illustrate their view of the AI design innovation space, the space where design thinking had the most impact (Figure 1).

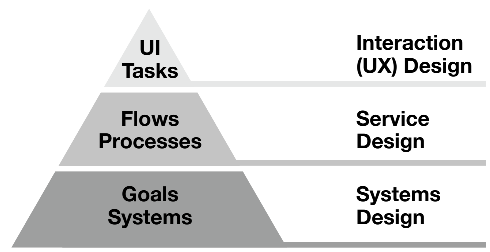

Figure 1: Three levels where designers can recognize or discover ways for AI to improve work. The width of each level indicates the value and impact design brings.

The diagram illustrates design activities at three levels: interaction (UX) design, service design, and systems design. At the top, AI optimized repetitive tasks that happen in an interface. At the bottom, AI helped to improve workers' performance, often by offering new insights or by augmenting their capabilities. Participants reflected on the tensions between automation and augmentation, and referred to the bottom level as a richer space for design: *"Often we hear the narrative "we'll automate the low value tasks for people to move to better jobs", but no one designs what those jobs are. [Designers can create value in cases where] you're not going to reach a 100% automation as the data itself is changing over time, and there is a role for the human in the loop to deal with the hard cases, but also to directly train the algorithm ... [Design can have a positive impact] for the users as they can employ their skills; the future of their job takes on some ownership for that model." (P11)*

Designers spoke about these three levels as a continuum, and they recognized new opportunities for AI across the space. They shared examples of recognizing opportunities at each level:

Interaction (UX) Design (UI, tasks): Designers had worked on a tool to classify financial transactions. This application divided operational work between the AI system and the human workers. The AI system automatically classified the most common and easily recognized transactions, and the workers classified the infrequent, uncommon transactions that required their expertise and common sense. As the interface design for the workers took shape, designers realized that the AI system could also classify many of the uncommon transactions correctly. This switch accelerated the pace of the work. Based on the new interaction design, workers only had to confirm that the classification was correct, or they would repair the misclassified transaction. *"The labels were there already. So we could use that as a placeholder to say 'is this right' rather than asking [workers] to fill it from scratch." (P5)*

Service Design (Flows, Processes): When designing an AI decision support tool for the pharmaceutical industry, designers conducted design research to understand scientists' mental models and workflows. They found out that scientists search various web sites for data regarding clinical trials. Realizing that this was crucial to their workflow, the team worked to ingest some of the data into the AI system: *"We circumvented linking off to a website that doesn't necessarily fit with their flow and pulled that data in and reorganized that in a way that suits them better." (P5)*

Systems Design (Goals, Systems): Designers had worked on an AI system to discover relevant relationships between medical diagnoses and treatments [87]. They created a tool where clinical experts (typically nurses) reviewed the discovered relationships in order to validate that this might be relevant. When a relationship was approved, it was forwarded to data scientists who used it to update the knowledge graph. One human-AI challenge was to motivate high quality work from the clinicians. The team reframed the role of these experts, shifting away from thinking of them as "coders" and explicitly referring to them as "AI curators". This shift served to *"upskill them, allowing them to be AI producers without becoming data scientists" (P12).* The interaction design used clinicians' expertise to build the AI's knowledge graph; clinicians directly trained and maintained their AI system. The new design simultaneously enhanced job satisfaction and improved AI learning and knowledge discovery.

Designers shared that the complexity of the AI system changed the level where design thinking could impact innovation. P11 shared that for simpler AI systems, *"...designers can focus on the UX, we can use traditional [UX] methods and the target of analysis can be the human user."* For more complex AI systems, *"...design initially brings more value by mapping the functional system, its goals, and its causal relationships. The methods move towards systems design, and the target of analysis is the socio-technical system."* The challenge of envisioning how humans and AI systems collaborate and the division of work between these two different types of intelligence created a richer space for designers to draw on their creative skills.

Recognizing an AI opportunity required three things. First, designers needed an internalized understanding of AI's capabilities, and they needed to notice the availability of data required for a specific capability. Second, they needed to conceptualize how the idea would lead to a co-creation of value between the user and the service. Third, their ideas needed to be viable, meaning that any AI feature requiring an additional cost would be assessed in terms of its value generation against the cost of development: *"It's always come from this joint realization of, this is doable within the budget constraints and access to data we have, but also this feature that we're asking the user to do is really boring and repetitive without it." (P5).*

When asked which actions helped them recognize AI opportunities, designers spoke about observing users, creating scenarios, participating in workshops, and wireframing. While conducting research, they would notice user behavior patterns and repetitive tasks. P9 shared that they spotted opportunities *"... whenever I start looking at different user types . . . repetitive tasks, processes, and routines."* They mentioned recognizing opportunities during ideation, when they sketched and generated scenarios. This happened both when working alone as well as in more structured group activities such as co-design workshops. They shared that wireframing offered one of the best moments for discovery: *"...when you put the input area on an interface. . . you could say there's an autocomplete here, or a suggestion box. That's the exact moment for me." (P4)* Below we detail the practices and approaches of participants. While some of these are specific to working with AI, some are applicable to more general design and data work (e.g. data-driven design).

4.1.1 AI Capability, Data, and the Data Pipeline. When starting a new project, designers invested significant time to understand the AI described in the design brief. They worked closely with data scientists, software engineers, and AI engineers to understand how the proposed AI system would work and the data required to make it work. P5 described this as, *"trying to understand the technicallness of what's going on with an AI solution ... what data is there that we can use as a material. Almost like treating data as you would a dropdown or other design material. What does the system know that we can then leverage for the UX?" (P5)* They established a shared understanding of the AI system.

Through a process of gaining an understanding of the intended AI system across several projects, designers developed a deeper understanding of AI's overall capabilities. This helped them learn to recognize new opportunities: *"The guise of us [designers] coming up to speed on what's going on in the project – as a consequence of that, we start to see those opportunities, saying, we know this already, so we can use that to drive some other feature." (P5)* Designers used several techniques to engage their technical collaborators and understand a system including diagraming, mapping logic flows, and data visualizations. These functioned as boundary objects between the design and data science expertise.

> *"A lot of our job as a design team on this project was sitting with data scientists and making them draw on a white board how it worked over and over again. ... I need to draw out a flow box diagram, or logic flows, and sit down and explain it with people until we get a shared understanding or mental model of what [the AI system] does." (P5)*
>
> *"We generated loads and loads of R plots to see what the outputs might look like, how you might classify those outputs. And then we could go back to the data scientists with these plots on the wall, and start asking them to annotate, label, and explain to us." (P8)*

Designers largely worked on the first versions of a new system, making access to user logs impossible. Data was frequently unavailable or difficult to access. As a workaround, they sometimes worked with a scheme describing the structure of data that might be available or they worked with training data being used to prototype the AI system. Even without ready access to data, designers invested great effort in exploring data and worker benefit as part of the system concept. *"When I'm doing data exploration, I look for hooks and what those hooks can mean in terms of functionality or interactions that you have in your UI. . . . for example if there's longitude or latitude, it suggests a map or some mapping functionality." (P8)* They frequently asked, *"What is the action you want to take from that dataset or what is the insight it's telling us?"* They frequently pushed

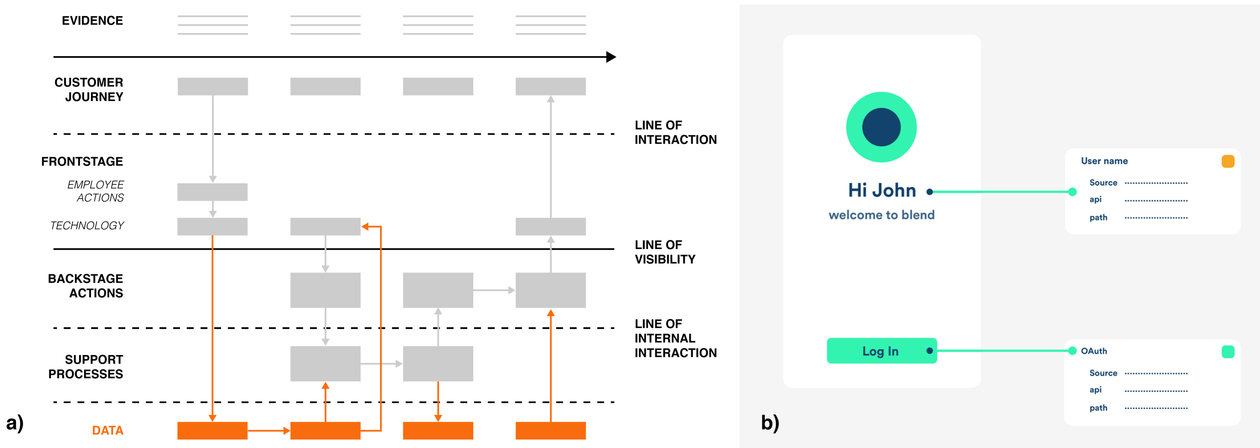

Figure 2: Augmented tools, such as (a) service blueprints with a data swim lane, and (b) annotated wireframes supported designers in understanding and communicating the role of data within their design.

clients and other stakeholders to gain access to data resources. They noted that the design and analytics teams often worked together to figure out if more data is needed: *"Say, there are six things we need to stitch together to find an answer. We know the three of those. We're still trying to get the other three attributes of the data." (P4)*

In addition to the data and its structure, designers stressed the importance of understanding the data pipeline: *"It's not just understanding what data we have available, but it's understanding which systems the data sits on, and whether data can be transferred across systems to be used together. So really the piping, can we pipe this data out of this system into this other system in order to achieve a particular goal?" (P10)* They often made system maps to learn the overall data flow between the front and back end. Diagramming and mapping helped reveal design opportunities beyond the interface (P5, P9, P10). While designers shared that combining datasets across sources may reveal new design possibilities, they raised several concerns around privacy and ethics (P1, P3, P9, P10, P11, P12). Participants were skeptical of the use of data and AI for adaptation or personalization in enterprise applications. They noted that features requiring user models may enable employers to infer worker's productivity, and could be instead designed through customization without the use of data or AI. While our focus was on envisionment, our discussions surfaced tensions around system boundaries – how much a system knows about its users versus how much it *actually needs* to know.

4.1.2 Tools, Methods, and Resources. Designers talked about their use of design tools and methods: *"Going from UX to service, it's a lot of the same [design] tools but you're expanding your reach." (P10)"* They used a combination of UX methods (e.g. interaction flows, personas, scenarios, user journeys), and service and systems design methods (e.g., service blueprints, systems mapping, causal loop diagrams). To address the challenges of working with AI and AI's need for data, they augmented service blueprints, adding data as a distinct swim lane (Figure 2a). They spoke about this as a way of visualizing the data pipeline. They annotated the data swim lane to describe the role data played, and they annotated wireframes to indicate the data source for specific UI features (Figure 2b).

In addition to augmenting service blueprints, they also tweaked the service model canvas, creating what they referred to as a "data-driven service design canvas" (Figure 3a) They frequently used this tool to support ideation and team alignment around data needs. Based on the canvas, they created a set of logic statements using the structure, *"if this, then that."* These statements aided ideation, exploration, and scenario construction: *"We give people post-its where they put [if, and, then] clauses together with actions, so 'if nothing was rejected on the last delivery, then repeat shopping list'." (P9)* The canvas explicitly prompted designers to think about the AI's value proposition and required data through questions such as *"how will this service help to make people's lives better?"*, *"when is the service triggered?"* and *"what data is needed at each point?"*. This exercise helped them build sophisticated data-driven services.

Drawing from data science's use of the terms insight, action, and outcome, designers created an exercise meant to capture the connections between these elements [54]. Data scientist collaborators were asked to complete cards that displayed the prompts: *"I want to know [Insight] so that I can [Action] to enable [Outcome]"* (Figure 3c). This exercise enabled the team to identify useful features and functions of the AI system: *"There were many different features that [the technical team] could begin to engineer but we were trying to figure out what was important to see, what was important to interact with. ... [this exercise] allowed us to spill out bite sized, finite pieces of ideas to then together formulate something coherent." (P8)*

To help improve their collective understanding of AI capabilities, designers created new resources [18, 19], such as the AI Creative Matrix (Figure 3b). This was meant to translate well known AI mechanisms such as NLP and computer vision into AI capabilities designers could envision from. The matrix used action verbs such as see, read, and hear, which made the capabilities explicit and put them in terms of human capability. They explained that the capabilities within the matrix also needed AI exemplars to become actionable to designers: *"One of the key things around our design and AI educational materials was examples. For a lot of people you*

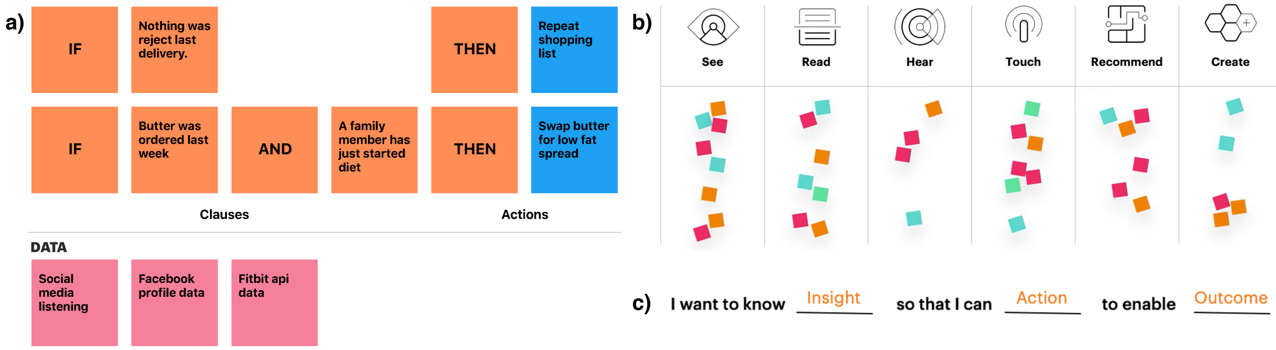

Figure 3: (a) Data-driven service design and systems design methods scaffolded designers' thinking around the AI system and its dependency on labelled data, (b) the AI Creativity Matrix was used to learn about and ideate AI capabilities, (c) Insights-Actions-Outcomes cards helped breaking down AI features into pieces for formulating ideas.

actually have to show, you need to give them those inspirational examples." (P11) For instance, when talking about *"seeing"* as a capability, they described a system that used computer vision and NLP. This system could *"see"* text on packaging. It then *"read"* the text it found, extracting the ingredients in order to monitor for a conflict with a known set of food allergies. They thought of these capabilities as functions that could be combined, such as *"seeing"* text and then *"reading"* any found text. Using action verbs instead of technical AI terms and mechanisms made ideation workshops more accessible for designers, product managers, clients, and other stakeholders that did not have AI or data science training.

4.2 Co-creating Value and Barriers to AI Innovation

Designers spoke frequently about value using service design language. For example, they talked about the co-creation of value between the company (service provider) and workers (users). The terms *"accelerators"* and *"enhancers"* were used to talk about the value of AI. AI accelerators speed up the pace or reduce the effort for current work, often fully automating input tasks. Enhancers improve the quality of output and the experience of work.

Proposals for a novel use for AI or a new form of human-AI interaction had to be proven as a business case to justify the investment. The value co-creation embodied in designers' concepts needed to easily outweigh the development and operational costs for building and deployment. Value for an accelerator was easier to quantify, estimate, and justify. Enhancers seemed more challenging to justify. Some value propositions designers used included increasing job satisfaction, enhancing decision making, improving the quality of data collection, and capturing organizational knowledge that might have potential future use. Designers shared that experiential value was often impossible to estimate without building experience prototypes: *"It's not just user acceptance testing. If you're actually measuring the [AI system's] impact, you have to simulate the thing that you want to measure." (P11).* They gave an example of an interactive decision support system where they had to build a simulator with a new set of metrics and key performance indicators (KPIs) to assess the value that the AI system might generate [82].

Innovations at the UI level, such as AUIs, most frequently took the form of accelerators that speeded work. Interestingly, while the value in terms of saved worker time was easier to estimate, designers described these innovations as a much harder to pitch as a convincing business case. *"If you start [innovating] from the UI, you need to convince the next levels as the business case provides the constraints." (P11)* Business constraints, including project timelines and budgets, played a critical role in determining what would and would not be included in a design. In most cases, the easily estimated value for a UI innovation was considered too low compared to possible development costs or risk to tight project timelines.

Designers used a simple heuristic to think about the cost of an AI feature: 1) The idea doable with the existing AI model and the dataset, 2) It requires collecting new data, 3) It requires building a new AI model and collecting new data. Consequently, the AI opportunities that designers searched for would often repurpose existing data or include only small extensions to the currently collected data. If an idea required AI development with additional cost, such as building a new learning model, it would likely move out of the current plan and get added to the future product roadmap. In some cases, the value for an innovation was not considered large enough, while in other cases, the value proved difficult to estimate. When deciding whether to pitch an idea, designers faced a challenge in that they did not have a good sense of the development effort for their idea in terms of cost or time.

Designers all agreed that estimating software costs was not their job. The opacity of this estimate seemed to discourage them from suggesting AI features. They tended to refer to any AI feature that required additional investment as "costly" or "expensive". We discussed several AI features in existing products and services to delineate what is cheaper, and what is more expensive. For example, building a single model for a system would be cheaper than building a separate model for each user. Similarly, building static models or models that are updated infrequently would be cheaper than building models that required constant data collection and frequent model building. Our conversations surfaced these cost-related AI properties as key aspects for designers to get a rough estimate of "how expensive" a proposed AI feature might be. However, little confidence was expressed in designers' ability to make an accurate estimate, with the exception of design leads. Designers had little understanding of the "expensiveness" of different AI capabilities and data remained more elusive.

In discussing an innovation's value and the challenge of making a business case, designers frequently described their work as defining an MVP – the minimum viable product. They described the culture of their work as very fast paced with strong budget and time constraints. This demanded a focus on articulating and refining a core set of features that co-created value. They used tools such as the impact effort matrix [33] to rapidly qualify and assess ideas: *"If we recognize that the feature has a very high value for the user and requires a very low [design and implementation] effort from the team, that's a quick win. If it is something that has some uncertain value to offer the user and it costs a lot, usually you tend to park that under the 'nice to have' or 'let's consider this in phase 2'." (P1)*

Designers rarely worked on phase 2 of a project. They rarely were informed about whether ideas saved for the future were followed up on. They drew a distinction between the type of design they did and the work that designers assigned to a specific product might do. They speculated that a designer assigned to a product was in a much better position to add in UI level accelerators that could enhance an already existing design. They noted that one major difference is that once a system was deployed, there would be user log data, which would provide clear evidence for how long it took to complete a task and how frequently that task was executed. Because they worked on first versions of a system or on redesigns, they rarely had access to this type of data. Prior work on UX designers with AI experience shows they would make use of this kind of data to envision innovations for existing systems [92].

In discussions about why they so infrequently innovate with AUIs, designers pointed out two additional, interrelated barriers. First, they shared that almost none of the enterprise systems they worked on employed a user model that held a detailed history of specific user interactions. While many online, customer-facing services employ user models to personalize a user's experience, enterprise systems instead mine worker roles from relatively smaller user bases (<10,000). Making a change that required data collection and model building for each worker seemed very expensive, and the performance gains seemed insufficient due the low volume and frequency of the tasks. So even when designers recognized opportunities where an AUI would help, they often held it back as the small amount of time savings would never outweigh the cost.

Second, they spoke about data availability as a barrier. Their projects often had uncertainty around data ownership, use, and privacy. This happened most frequently when they worked on platforms owned by parent companies and operated by client businesses. It was difficult for teams to have access to data for any reason outside of core functionality. Designers frequently asked *"is there data available?"*, *"who owns the data?"*, and *"are we allowed to use it for the purpose that we want to use it for?"* Working with clients' transfer teams meant that they had no access to the data needed to design and refine the interaction around AI models after deployment, something any system using personalization would need. Additionally, enterprise organizations had to clearly state the purpose of data collection and use in client contracts upfront, due to IP generation as well as privacy regulations, such as GDPR compliance. Overall, the challenges around data access and use posed constraints on data exploration and the search for emergent value in the data.

4.3 Collaboration in AI Teams

We wanted to better understand the collaboration between designers and data scientists. We asked participants what practices they used to overcome the gaps in collaboration. Designers emphasized co-located, informal collaboration in successfully spanning the cross-disciplinary gap. Below, we describe the role of designers in AI teams, and then discuss collaboration and the role of boundary objects in the work.

4.3.1 Role of Designers in Cross-functional AI Teams. Design practitioners supported AI teams in three ways: 1) designing the human-AI interaction, 2) facilitating alignment, and 3) broadening AI's value space. The first two activities were part of all projects, while the third seemed to happen less frequently on select projects.

Designing Human-AI Interaction: A principal task for designers was to design human-AI interaction. Designers typically joined ongoing projects, after the data science team performed analytics and developed a proof of concept algorithm for a particular use case. Joining a team meant that at this stage, designers worked on a predefined AI problem defined by the clients and shaped by the team: *"There may be a very defined business area where the client already has data and they know what the value is." (P11)*

The design process for designing human-AI interaction did not differ radically from traditional design processes. However, designers spent more time in the early stages of alignment to be able to *"frame what it is [the technical teams and the client] want to do." (P9)* Designing human-AI interaction required special attention to usability issues and user acceptance. An essential part of designers' work was situating AI in users' workflow: *"If we're designing an AI solution that's going to take away part of [users'] work or responsibility, we really need to understand how they think about that so that they can trust the outcome, but also it fits better to their workflow." (P5)*. Participants highlighted the importance of design research to understand user needs around trust, explainability, interpretability, transparency and acceptance of AI systems.

Designers spoke of "design as communication": *"We usually work with the outputs [of the model]. We innovate by improving flows, interaction concepts or improving the adoption of UI components." (P1)* This often involved presenting the complex output of AI systems to end users in ways that supported their mental models. For example, when designing a contract risk analysis tool, designers worked on representing the level of risk for a given contract. Working with the business and data science teams, they decided on confidence score thresholds to rank and surface riskier contracts for users to immediately act on. Designers often thought about error recovery and potential errors users can introduce when generating labels. They frequently asked data scientists *"What would happen if the user did this, is that going to ruin the algorithm?" (P5)*. Some projects required making speed and accuracy trade-offs, and an understanding of users' willingness to wait and tolerate delays. In such cases, UX design had a direct impact on algorithm selection (P1, P11).

Facilitating alignment: Similar to research describing designers' role as facilitators [58], participants often spoke of themselves as facilitators on AI teams. They played a key role in alignment between disciplines: *"The designer is someone who has to go between data scientists and software engineering because design speaks visual language and everyone can look at it and point at things." (P5)* In initial phases, designers worked to facilitate stakeholder alignment in terms of setting project goals, requirements, and success metrics. They used knowledge elicitation exercises and boundary objects [section 4.3.2] to help technical teams and clients articulate their AI-driven objectives. In later development stages, they held workshops with stakeholders to discuss and prioritize features through scenarios and wireframes: *"That's where we get together all ideas and make the plan for what the product or prototype should be and set those functional requirements." (P4)*

Broadening AI's value space: Depending on the project, participants occasionally engaged in problem setting – envisioning and reframing to align on the right design. As described in early stage AI strategy projects: *"We have clients who come in for an open innovation session where they have some key strategic areas they want to go. Part of the goal is helping them ideate about potential applications of AI using design thinking methods. If you get [clients] at that stage, you don't know if they have the data [for an AI application]." (P11)* In these cases, participants mapped out the problem space using concept mapping, then worked to identify models and data types that can drive particular design goals. They held co-creation workshops with all teams and stakeholders, and engaged their clients in idea generation using AI ideation tools, such as the AI Creative Matrix. This approach to envisioning AI strategy seemed more similar to agile development process than traditional user-centered design.

There were also cases where designers were able to broaden the project framing using design thinking and design research to get to the root causes of problems. One example was a project in the public safety domain: *"The primary objectives given by the client would only solve a subset of issues. So how do you then also solve the periphery issues that are associated with making this more effective?" (P3)* They used systems design methods including systems mapping and causal loop diagramming to explore the relationships between technical, cultural and organizational challenges, leading to discovery of emergent user and business value (P3, P4, P11).

4.3.2 Collaboration and Boundary Objects. Co-located work was an intentional organizational decision that made design and data science collaboration easier. Data scientists stated that by working with designers, they became more conscious of their assumptions about end users' understanding of AI system outputs: *"As a data scientist, I bring a lot of assumptions, 'of course [the users] are going to understand this' ... [I realized] that we need to take more time thinking about the outputs with the designers." (P12, P13).* It was noted that collaborating with designers early in the process leads to *"a happier marriage" (P13).* One data scientist shared that in addition to increasing business value, they became aware of the "experiential value" of the felt experience of users: *"[The design] allowed [users] to share their expertise. This was an experience that they really value now in their new role. ... I wouldn't have been sensitized to the potential value on that experience side if it weren't for the designers on the team." (P14).* The data science team also benefited from working with designers in eliciting requirements from domain experts, as design practice has well-established methods and tools for knowledge elicitation from end users (P10, P11, P12).

Design practitioners shared that working with data scientists made them more data aware. They also seemed to be running their ideas by data scientists, using their expertise as a proxy [94] for understanding AI's design possibility space:

> *"As a designer, I'm not often sure of what's required to run a particular idea I have. That's sort of the nature of the conversation, which is I draw something. I'll go to someone on the team and say, can I talk to you about this? I'm thinking it's going to be useful, I think it's possible, but I actually don't know. So I need your input. Hopefully they come with a constructive attitude to build on the idea." (P5)*

Design practitioners stated that working with data scientists helped them to identify the root cause of pain points, especially in the research phase. They shared including all team members in design research as a best practice, especially with data scientists: *"A data scientist will want to know something very particular, like 'do [users] use this kind of data?' They get an opportunity to reprobe certain areas that we might not have probed as we don't have the experience or knowledge." (P10)* They noted that this collective approach to design research yields more value (P2, P5, P8, P10).

Close collaboration was not without challenges. One designer shared that the lack of a common language was a challenge: *"Even though [the data scientists] are speaking English, it's like they have a different language. What does it mean to have outputs from a knowledge graph in practical terms?" (P9)* We discussed artifacts that facilitated communication and collaboration across roles throughout the AI development process. Participants readily identified several boundary objects, including whiteboard sketches and visualizations, annotated wireframes, and service blueprints with data annotations:

> *"When that data layer was added [to service blueprints], it made a huge difference in terms of the data scientists being able to talk through the process with the designers. Because it's really important for us where the data feeds in, so we know when we can use it for our analytics and AI." (P12)*
> *"[annotated wireframes] was designed to make our conversations with the development team a lot easier, because you're drawing this box, but where does this box pull its data from?" (P5)*

These responses provided evidence of the use of boundary objects for scaffolding conversations between design and data science. Boundary objects were used in sketching to facilitate a shared understanding, and in prototyping to detail data dependencies.

5 DISCUSSION

Prior research has shown that UX practitioners face difficulties in envisioning new AI products and services and collaborating with data scientists [21]. Recent work suggests AUIs might be one type of AI innovation designers could champion in the enterprise. We set out to discover why AUIs were not being used; however, over the course of this work, our focus broadened to explore how design practitioners working on enterprise projects effectively innovate with AI. We found that design practitioners recognize and recommend a broad range of AI innovations. The opacity around an AI feature's cost and the sense that the cost will outweigh the value for an innovation created a real barrier to design innovation with AI. It caused designers to hold back their ideas. This work also showed that designers play various roles on cross-functional AI teams, and that AI teams benefit from close collaborations. We reflect on these findings, their implications for research and practice, and opportunities for development of new tools that can support designers in effectively working with AI and collaborating on AI teams.

5.1 Implications for Research

Design is a reflective practice where designers engage in a "conversation with materials" to envision things that do not yet exist[73]. Materials "talk back" to designers, revealing possibilities and constraints. HCI research has explored the idea of AI as a design material, framing AI as a set of technical capabilities designers can use to create novel features, products, and services [21, 36, 93]. By reflecting on the insights that our study revealed around how experienced designers engage data and AI, we deepen the concept of AI as a design material and suggest areas for future research.

5.1.1 AI as Design Material. Previous research investigating AI as a design material has shown that designers who had built a large set of "designerly abstractions" of AI's capabilities were more successful and comfortable at working with AI [92]. These designers were able to engage in reflective conversations with AI, and they leveraged data scientists' technical expertise as proxy for material "talk back" on what is possible [64, 92].

Our study confirmed this. Participants had an implicit understanding of AI's capabilities and they frequently envisioned and recognized opportunities where AI could add value. This "good enough" understanding of capabilities, combined with close collaboration with data scientists, enabled designers to have reflective conversations with AI as a design material. In addition to a general understanding of AI capabilities, designers worked to gain an in-situ understanding of the AI system – what the AI system knows and what it is doing in that particular context with that particular data. Designers who wish to innovate with AI will need to prepare themselves for noticing the availability of a dataset and exploring it; envisioning the data pipeline; and effectively communicating how value co-creation is likely to generate sufficient, measurable impact. Design researchers could create resources that document these design-specific material aspects of AI. These efforts have the potential to advance the HCI community's understanding of what an adequate AI literacy means for design practice.

5.1.2 Design Tools and Methods. Our study provided details on the types of tools, methods, and exercises that helped to envision novel forms and functions of AI. Designers thought of AI capabilities as action verbs implying a human function (e.g. read, see, listen) as opposed to thinking about the technical mechanism (e.g., neural networks, collaborative filtering). We suspect that the AI capabilities designers identified would generalize to many contexts and applications beyond the enterprise. The self-made tools participants developed for internal use, such as the AI Creative Matrix (Figure 3b), suggest a need for new tools to help with envisioning. Taxonomies and resources can be created that explicitly document AI capabilities with exemplars to help designers operationalize AI concepts. Our study showed that non-experts also benefit from these resources in participating collective AI ideation. Additional research is needed to explore the potential forms and representations that can situate these tools and artifacts in current design practice and process.

Participants also used diagrams, data visualizations, and service/systems design tools such as system maps and the service design canvas to elicit knowledge from data scientists about the AI system, the dataset, and the data pipeline. These tools served as boundary objects between design and data science expertise; they helped in discussing the data dependencies, in identifying root causes, and in formulating novel and coherent AI system behaviors.

Designers would benefit from new knowledge elicitation tools and exercises that support team alignment on AI systems. These tools can potentially surface AI's design opportunities as a byproduct of the alignment process. Recent research has investigated the use of data-augmented tools, especially for service design [50]. HCI researchers should evaluate and improve these sketching and prototyping tools, building on tools augmented with data, such as annotated wireframes and service blueprints with a data swim lane.

5.1.3 AI in Enterprise vs Consumer Space. Previous HCI research has largely characterized design-led AI innovation as personalization of consumer facing products and services [92]. It has also identified user behavior or telemetry data as the data type designers work with. In the context of enterprise software, personal data, such as user logs, often was not available due to working on the first version of a new system or due to privacy concerns. Datasets instead took the form of training datasets, such as knowledge graphs, and mockup data as an approximation of the target data scheme. Additionally, enterprise applications had much smaller user bases, which made it less likely for AI concepts to bring large enough value to move into products. Enterprise designers conceptualized value as it relates to users and business, including "accelerators" that optimize tasks, and "enhancers" that improve the quality of human decision making or worker experience. Design researchers should consider the diverse ways professional practices are situated within companies across contexts, and how design relates to the conceptualization and monetization of different types of value.

5.2 Implications for Collaborative Practices

This study revealed several research opportunities for supporting cross-functional collaboration in AI teams through: 1) co-located design and data science teams, and 2) boundary objects.

5.2.1 Co-located Design and Data Science Teams. Prior work reports a gap between design and data science practice: designers envision AI innovations that cannot be built, and data scientists propose AI innovations that users do not want [91]. The design and data science teams in this study were able to span this gap through co-located, informal collaboration. Data scientists helped designers assess the technical feasibility of their ideas. Designers supported data scientists in working with end users for knowledge elicitation. Future research should investigate the benefits and challenges of emergent collaborative practices on industry AI teams with an eye for how these roles can complement each other. This finding also presents implications for HCI education: programs preparing design students for industry should consider providing opportunities for working with students from data science programs.

This study showed that although occasional, designers do engage in reframing "the right AI thing to design". Participants' roles expanded beyond designing human-AI interactions to ideate on the value AI could and should bring. Prior research characterized problem setting and reframing as designers' forte [24, 96], yet noted that AI-driven interactions are difficult to abstract and reframe [93].

This finding about AI problem setting provides an opportunity for further research. What roles can and should designers play in cross-functional AI teams? What types of value and impact could designers bring when they are involved in designing the right AI thing? Our study showed that designers leveraged their role as facilitators to involve end users in AI development processes. Recent research has highlighted the need for a principled discussion on broadening stakeholder participation in AI through design to account for inclusiveness or fairness goals [17, 75]. More research is needed to define designers' role in AI design and development.

5.2.2 Boundary Objects. Recent HCI research has highlighted the use of boundary objects to scaffold collaboration across different roles and disciplines in industry AI teams [14, 37, 43, 91]. Our study showed that boundary objects are critical in facilitating an effective collaboration between designers and data scientists. Artifacts including flow diagrams, system maps, and service data blueprints supported participants both in envisioning to establish a shared understanding, and in prototyping to detail the data dependencies. We see an opportunity for developing and assessing boundary objects. What new boundary objects might help AI teams in AI problem formulation? Could these boundary objects be augmented to scaffold discussions around fairness, bias, and privacy?

In parallel, more investigation of how collaboration unfolds across multiple roles in AI teams is also needed. Our study focused on designers and data scientists, yet participants frequently mentioned other roles, including business managers and software developers. What type of boundary objects might help bridging multiple disciplines and stakeholders throughout the AI lifecycle? More research is needed to investigate how boundary objects might facilitate collaboration in various AI development contexts.

5.3 Open Research Questions

Our study raised several open research questions that merit further study. Below, we detail two challenges for future investigation.

How to Empower Designers to Guesstimate AI's Cost and Value? Our study revealed an opacity around the cost of proposed AI features. This acted as a filter: participants were less likely to propose AI features that required an additional cost. While it is not designers' job to estimate development and operational costs to continuously collect data and rebuild models, designers can benefit from a high level understanding of how expensive an idea would be. For example, proposing to build a static model for a group of users will likely be cheaper than building a dynamic model for individual users. Another possible solution could be sensitizing designers to low hanging fruit – relatively less costly AI solutions or off-the-shelf AI libraries and pre-built models that do not bring an additional cost.

Similarly, designers would benefit from detailing and assessing the value proposition of an AI system. Recently, industry researchers have shared methods for measuring UX in the early phases of AI product development to rapidly assess and prioritize AI features [74]. Our study echoes this. The ability to effectively qualify and communicate how value co-creation is likely to outweigh costs could support designers in recommending and prioritizing AI opportunities with their teams.

What Lenses Could Help in Designing for AI? HCI research has largely focused on the needs of UX designers when investigating the implications of AI on design practice. This case suggests that service/system designers –people that do early, strategic design work– should also be a focus as HCI investigates human-AI interaction. Recent research has explored design processes beyond user-centered design, such as service design [27, 28, 42, 72] and more-than-human design [29, 62] for designing for data and AI. Others have identified systems thinking and cybernetics as useful frameworks for designing complex AI systems [23, 60]. Design and HCI researchers could explore these additional lenses to better define and broaden AI's design innovation space.

6 LIMITATIONS

Our study had three limitations. First, the study is exclusively from the perspective of a small group of participants who are highly specialized in working with AI. We do not know if practitioners in other enterprise organizations have similar experiences or collaborative practices. Second, our team had a specific focus on the practices of design and data science practitioners on AI teams. We note that there are several other practitioners and stakeholders that merit further study. Finally, while our study raised several concerns around privacy and ethics around data use, our focus was on envisioning with data and AI capabilities. We note that issues around AI ethics, safety, reliability, fairness, transparency, and accountability are central to the HCI community's research focus and requires further study from practitioners' perspectives [6, 15, 25, 37, 49, 56, 85]. We invite researchers to investigate the practices of AI teams around responsible AI to further scaffold cross-disciplinary collaboration in mitigating potential harms.

7 CONCLUSION

This paper presented a study that explored how experienced designers on AI teams in the enterprise work and innovate with AI as a design material. We expanded previous research by providing a rare description of how and when design practitioners envision and identify AI opportunities, and the roles they play on AI teams. We encourage HCI researchers to join us in exploring and redefining the role design can play in the creation of AI technologies that benefit people and society.

ACKNOWLEDGMENTS

This work was partially supported by the National Science Foundation under Grant Numbers 2007501 and 2112633, and by Accenture. We thank Matias Ferrero, Gemma Gallagher, Antonio Pacheco, and Navdeep Sharma for their valuable input in our initial discussions.

REFERENCES

[1] Accenture. 2020. Accenture Logistics Platform: Win the last mile. Retrieved Mar 24, 2022 from https://www.accenture.com/gb-en/services/public-service/accenture-logistics-platform

[2] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N Bennett, Kori Inkpen, et al. 2019. Guidelines for human-AI interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–13.

[3] Apple. 2019. Human Interface Guidelines: Machine Learning. https://developer.apple.com/design/human-interface-guidelines/machine-learning/overview/introduction/

[4] Jesse Josua Benjamin, Arne Berger, Nick Merrill, and James Pierce. 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*. 1–14.

[5] Reuben Binns, Max Van Kleek, Michael Veale, Ulrik Lyngs, Jun Zhao, and Nigel Shadbolt. 2018. 'It's Reducing a Human Being to a Percentage' Perceptions of Justice in Algorithmic Decisions. In *Proceedings of the 2018 Chi conference on human factors in computing systems*. 1–14.

[6] Abeba Birhane. 2021. Algorithmic injustice: a relational ethics approach. *Patterns* 2, 2 (2021), 100205.

[7] Sara Bly and Elizabeth F Churchill. 1999. Design through matchmaking: technology in search of users. *interactions* 6, 2 (1999), 23–31.

[8] Kirsten Boehner and Carl DiSalvo. 2016. Data, design and civics: An exploratory study of civic tech. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2970–2981.

[9] Sander Bogers, Joep Frens, Janne Van Kollenburg, Eva Deckers, and Caroline Hummels. 2016. Connected baby bottle: A design case study towards a framework for data-enabled design. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. 301–311.

[10] Nico Brand, William Odom, and Samuel Barnett. 2021. A Design Inquiry into Introspective AI: Surfacing Opportunities, Issues, and Paradoxes. In *Designing Interactive Systems Conference 2021*. 1603–1618.

[11] Virginia Braun and Victoria Clarke. 2012. Thematic analysis. (2012).

[12] Carrie J Cai, Samantha Winter, David Steiner, Lauren Wilcox, and Michael Terry. 2021. Onboarding Materials as Cross-functional Boundary Objects for Developing AI Assistants. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–7.

[13] Stephanie Chamberlain, Helen Sharp, and Neil Maiden. 2006. Towards a framework for integrating agile development and user-centred design. In *International Conference on Extreme Programming and Agile Processes in Software Engineering*. Springer, 143–153.

[14] Ed Chi. 2020. A twist on loss functions as boundary objects. Retrieved Mar 24, 2022 from https://www.youtube.com/watch?v=jHTL_ysgetE&t=758s/

[15] Shruthi Sai Chivukula, Chris Rhys Watkins, Rhea Manocha, Jingle Chen, and Colin M Gray. 2020. Dimensions of UX Practice that Shape Ethical Awareness. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[16] Design Council. 2005. The 'double diamond'design process model. *Design Council* (2005).

[17] Fernando Delgado, Stephen Yang, Michael Madaio, and Qian Yang. [n.d.]. Stakeholder Participation in AI: Beyond "Add Diverse Stakeholders and Stir". *communities* 46, 34 ([n. d.]), 27.

[18] Accenture The Dock. 2019. Man with machine: designing a new future for humans and AI. Retrieved Mar 24, 2022 from https://medium.com/accenture-the-dock/man-with-machine-designing-a-new-future-for-humans-and-ai-4f43e9a94517

[19] Accenture The Dock. 2020. Designed Intelligence. Retrieved Mar 24, 2022 from https://youtu.be/pkFOcgD8AnA?t=551

[20] Graham Dove and Anne-Laure Fayard. 2020. Monsters, metaphors, and machine learning. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–17.

[21] Graham Dove, Kim Halskov, Jodi Forlizzi, and John Zimmerman. 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*. 278–288.

[22] Graham Dove and Sara Jones. 2014. Using data to stimulate creative thinking in the design of new products and services. In *Proceedings of the 2014 conference on Designing interactive systems*. 443–452.

[23] Hugh Dubberly and Paul Pangaro. 2019. Cybernetics and design: Conversations for action. In *Design Cybernetics*. Springer, 85–99.

[24] Daniel Fallman. 2003. Design-oriented human-computer interaction. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 225–232.

[25] Sina Fazelpour and Zachary C Lipton. 2020. Algorithmic fairness from a non-ideal perspective. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. 57–63.

[26] Rebecca Fiebrink, Dan Trueman, and Perry R Cook. 2009. A Meta-Instrument for Interactive, On-the-Fly Machine Learning.. In *NIME*. 280–285.

[27] Jodi Forlizzi. 2018. Moving beyond user-centered design. *interactions* 25, 5 (2018), 22–23.

[28] Andrea Gasparini, Ahmed Abdi Mohammed, and Gabriele Oropallo. 2018. Service design for artificial intelligence. In *Linköping Electronic Conference Proceedings*. Linköping University Electronic Press, 1064–1073.

[29] Elisa Giaccardi and Johan Redström. 2020. Technology and more-than-human design. *Design Issues* 36, 4 (2020), 33–44.

[30] Fabien Girardin and Neal Lathia. 2017. When user experience designers partner with data scientists. In *2017 AAAI Spring Symposium Series*.

[31] Google. Accessed September 2021. Material Design Machine Learning Patterns. https://material.io/design/machine-learning/understanding-ml-patterns.html#

[32] Katerina Gorkovenko, Daniel J Burnett, James K Thorp, Daniel Richards, and Dave Murray-Rust. 2020. Exploring the future of data-driven product design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[33] Nielsen Norman Group. [n.d.]. *Using Prioritization Matrices to Inform UX Decisions*. Retrieved Mar 24, 2022 from https://www.nngroup.com/articles/prioritization-matrices/

[34] Patrick Hebron. 2016. Machine learning for designers. (2016).

[35] Karey Helms. 2019. Do you have to pee? A design space for intimate and somatic data. In *Proceedings of the 2019 on Designing Interactive Systems Conference*. 1209–1222.

[36] Lars Erik Holmquist. 2017. Intelligence on tap: artificial intelligence as a new design material. *interactions* 24, 4 (2017), 28–33.

[37] Kenneth Holstein, Jennifer Wortman Vaughan, Hal Daumé III, Miro Dudik, and Hanna Wallach. 2019. Improving fairness in machine learning systems: What do industry practitioners need?. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–16.

[38] Matthew K Hong, Adam Fourney, Derek DeBellis, and Saleema Amershi. 2021. Planning for Natural Language Failures with the AI Playbook. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–11.

[39] Anthony Jameson. 2008. Adaptive interfaces and agents. *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications* (2008), 305–330.

[40] Xiaoneng Jin, Mark Evans, Hua Dong, and Anqi Yao. 2021. Design heuristics for artificial intelligence: inspirational design stimuli for supporting UX designers in generating AI-powered ideas. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–8.

[41] Bonnie E John, Len Bass, Rick Kazman, and Eugene Chen. 2004. Identifying gaps between HCI, software engineering, and design, and boundary objects to bridge them. In *CHI'04 extended abstracts on Human factors in computing systems*. 1723–1724.

[42] Titta Jylkäs, M Aijälä, Tytti Vuorikari, and Vésaal Rajab. 2018. AI Assistants as Non-human Actors in Service Design. In *In Proc of DMI Acad Des Manag Conf*. 1–2.

[43] Claire Kayacik, Sherol Chen, Signe Noerly, Jess Holbrook, Adam Roberts, and Douglas Eck. 2019. Identifying the intersections: User experience+ research scientist collaboration in a generative machine learning interface. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–8.

[44] R. King, E.F. Churchill, and C. Tan. 2017. *Designing with Data: Improving the User Experience with A/B Testing*. O'Reilly Media. https://books.google.com/books?id=eM-PDgAAQBAJ

[45] Rafal Kocielnik, Saleema Amershi, and Paul N Bennett. 2019. Will you accept an imperfect ai? exploring designs for adjusting end-user expectations of ai systems. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.

[46] Peter Kun, Ingrid Mulder, and Gerd Kortuem. 2018. Design enquiry through data: appropriating a data science workflow for the design process. In *Proceedings of the 32nd International BCS Human Computer Interaction Conference 32*. 1–12.

[47] Charlotte P Lee. 2007. Boundary negotiating artifacts: Unbinding the routine of boundary objects and embracing chaos in collaborative work. *Computer Supported Cooperative Work (CSCW)* 16, 3 (2007), 307–339.

[48] Min Kyung Lee. 2018. Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society* 5, 1 (2018), 2053951718756684.

[49] Min Kyung Lee, Nina Grgić-Hlača, Michael Carl Tschantz, Reuben Binns, Adrian Weller, Michelle Carney, and Kori Inkpen. 2020. Human-centered approaches to fair and responsible AI. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.

[50] Fan Li and Yuan Lu. 2021. Engaging end users in an ai-enabled smart service design-the application of the smart service blueprint scape (SSBS) framework. *Proceedings of the Design Society* 1 (2021), 1363–1372.

[51] Q Vera Liao, Daniel Gruen, and Sarah Miller. 2020. Questioning the AI: informing design practices for explainable AI user experiences. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–15.

[52] Joseph Lindley, Haider Ali Akmal, Franziska Pilling, and Paul Coulton. 2020. Researching AI Legibility through Design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[53] Daria Loi, Thomas Lodato, Christine T Wolf, Raphael Arar, and Jeanette Blomberg. 2018. PD manifesto for AI futures. In *Proceedings of the 15th Participatory Design Conference: Short Papers, Situated Actions, Workshops and Tutorial-Volume 2*. 1–4.

[54] Eoin Ó Loideáin. 2019. Towards algorithmic equality. Retrieved Mar 24, 2022 from https://medium.com/design-voices/towards-algorithmic-equality-118e6ba10b74

[55] Duri Long and Brian Magerko. 2020. What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–16.

[56] Michael A Madaio, Luke Stark, Jennifer Wortman Vaughan, and Hanna Wallach. 2020. Co-designing checklists to understand organizational challenges and opportunities around fairness in ai. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[57] Nirav Malsattar, Tomo Kihara, and Elisa Giaccardi. 2019. Designing and Prototyping from the Perspective of AI in the Wild. In *Proceedings of the 2019 on Designing Interactive Systems Conference*. 1083–1088.

[58] Ezio Manzini. 2009. New design knowledge. *Design studies* 30, 1 (2009), 4–12.

[59] Yaoli Mao, Dakuo Wang, Michael Muller, Kush R Varshney, Ioana Baldini, Casey Dugan, and Aleksandra Mojsilović. 2019. How data scientistswork together with domain experts in scientific collaborations: To find the right answer or to ask the right question? *Proceedings of the ACM on Human-Computer Interaction* 3, GROUP (2019), 1–23.

[60] Nikolas Martelaro and Wendy Ju. 2018. Cybernetics and the design of the user experience of AI systems. *Interactions* 25, 6 (2018), 38–41.

[61] Christine Murad, Cosmin Munteanu, Leigh Clark, and Benjamin R Cowan. 2018. Design guidelines for hands-free speech interaction. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. 269–276.

[62] Iohanna Nicenboim, Elisa Giaccardi, Marie Louise Juul Søndergaard, Anuradha Venugopal Reddy, Yolande Strengers, James Pierce, and Johan Redström. 2020. More-than-human design and AI: In conversation with agents. In *Companion Publication of the 2020 ACM Designing Interactive Systems Conference*. 397–400.

[63] William Odom and Tijs Duel. 2018. On the design of OLO Radio: Investigating metadata as a design material. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–9.

[64] Fatih Kursat Ozenc, Miso Kim, John Zimmerman, Stephen Oney, and Brad Myers. 2010. How to support designers in getting hold of the immaterial material of software. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2513–2522.

[65] Google PAIR. 2019. People + AI Guidebook. pair.withgoogle.com/guidebook

[66] Samir Passi and Steven J Jackson. 2018. Trust in data science: Collaboration, translation, and accountability in corporate data science projects. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–28.

[67] James Pierce, Phoebe Sengers, Tad Hirsch, Tom Jenkins, William Gaver, and Carl DiSalvo. 2015. Expanding and refining design and criticality in HCI. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2083–2092.

[68] Dominik Pins, Timo Jakobi, Alexander Boden, Fatemeh Alizadeh, and Volker Wulf. 2021. Alexa, We Need to Talk: A Data Literacy Approach on Voice Assistants. In *Designing Interactive Systems Conference 2021*. 495–507.

[69] David Piorkowski, Soya Park, April Yi Wang, Dakuo Wang, Michael Muller, and Felix Portnoy. 2021. How ai developers overcome communication challenges in a multidisciplinary team: A case study. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–25.

[70] Holger Rhinow, Eva Köppen, and Christoph Meinel. 2012. Design prototypes as boundary objects in innovation processes. (2012).

[71] Daniela K Rosner, Saba Kawas, Wenqi Li, Nicole Tilly, and Yi-Chen Sung. 2016. Out of time, out of place: Reflections on design workshops as a research method. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. 1131–1141.

[72] Virpi Roto, Jung-Joo Lee, Effie Lai-Chong Law, and John Zimmerman. 2021. The Overlaps and Boundaries Between Service Design and User Experience Design. In *Designing Interactive Systems Conference 2021*. 1915–1926.

[73] Donald Schön and John Bennett. 1996. Reflective conversation with materials. In *Bringing design to software*. 171–189.

[74] Julie Anne Séguin, Alec Scharff, and Kyle Pedersen. 2019. Triptech: a method for evaluating early design concepts. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–8.

[75] M Sloane, E Moss, O Awomolo, and L Forlano. 2020. Participation is not a Design Fix for Machine Learning. *Participatory Approaches to Machine Learning* (2020).

[76] Marie Louise Juul Søndergaard and Lone Koefoed Hansen. 2018. Intimate futures: Staying with the trouble of digital personal assistants through design fiction. In *Proceedings of the 2018 Designing Interactive Systems Conference*. 869–880.

[77] Chris Speed and Jon Oberlander. 2016. Designing from, with and by Data: Introducing the ablative framework. In *Proceedings of the International Design Research Society Conference*.

[78] Aaron Springer, Jean Garcia-Gathright, and Henriette Cramer. 2018. Assessing and Addressing Algorithmic Bias-But Before We Get There.... In *2018 AAAI Spring Symposium Series*.

[79] Susan Leigh Star and James R Griesemer. 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.

[80] S Stumpf, L Strappelli, S Ahmed, Y Nakao, A Naseer, GD Gamba, and D Regoli. 2021. Design Methods for Artificial Intelligence Fairness and Transparency. In *CEUR Workshop Proceedings*, Vol. 2903.

[81] Hariharan Subramonyam, Colleen Seifert, and Eytan Adar. 2021. Towards A Process Model for Co-Creating AI Experiences *(DIS '21)*. Association for Computing Machinery, New York, NY, USA, 1529–1543. https://doi.org/10.1145/3461778.3462012

[82] Connor W Upton and Fergus R Quilligan. 2014. GREYBOX scheduling: Designing a joint cognitive system for sustainable manufacturing. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. 897–900.

[83] Philip van Allen. 2018. Prototyping ways of prototyping AI. *Interactions* 25, 6 (2018), 46–51.

[84] Michael Veale, Max Van Kleek, and Reuben Binns. 2018. Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making. In *Proceedings of the 2018 chi conference on human factors in computing systems*. 1–14.

[85] Susann Wagenknecht, Min Lee, Caitlin Lustig, Jacki O'Neill, and Himanshu Zade. 2016. Algorithms at work: empirical diversity, analytic vocabularies, design implications. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*. 536–543.

[86] Greg Walsh and Eric Wronsky. 2019. AI+ co-design: developing a novel computer-supported approach to inclusive design. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*. 408–412.

[87] H. James Wilson and Paul R. Daugherty. 2019. Creating the Symbiotic AI Workforce of the Future. Retrieved Mar 24, 2022 from https://sloanreview.mit.edu/article/creating-the-symbiotic-ai-workforce-of-the-future/

[88] Jason Shun Wong. 2018. Design and fiction: imagining civic AI. *Interactions* 25, 6 (2018), 42–45.

[89] Allison Woodruff, Sarah E Fox, Steven Rousso-Schindler, and Jeffrey Warshaw. 2018. A qualitative exploration of perceptions of algorithmic fairness. In *Proceedings of the 2018 chi conference on human factors in computing systems*. 1–14.

[90] Qian Yang, Nikola Banovic, and John Zimmerman. 2018. Mapping machine learning advances from hci research to reveal starting places for design innovation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–11.

[91] Qian Yang, Justin Cranshaw, Saleema Amershi, Shamsi T Iqbal, and Jaime Teevan. 2019. Sketching nlp: A case study of exploring the right things to design with language intelligence. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[92] Qian Yang, Alex Scuito, John Zimmerman, Jodi Forlizzi, and Aaron Steinfeld. 2018. Investigating how experienced UX designers effectively work with machine learning. In *Proceedings of the 2018 Designing Interactive Systems Conference*. 585–596.

[93] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 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*. 1–13.

[94] Qian Yang, John Zimmerman, Aaron Steinfeld, and Anthony Tomasic. 2016. Planning adaptive mobile experiences when wireframing. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. 565–576.

[95] Amy X Zhang, Michael Muller, and Dakuo Wang. 2020. How do data science workers collaborate? roles, workflows, and tools. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (2020), 1–23.

[96] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 493–502.

[97] John Zimmerman, Changhoon Oh, Nur Yildirim, Alex Kass, Teresa Tung, and Jodi Forlizzi. 2020. UX designers pushing AI in the enterprise: a case for adaptive UIs. *Interactions* 28, 1 (2020), 72–77.

[98] John Zimmerman, Anthony Tomasic, Charles Garrod, Daisy Yoo, Chaya Hiruncharoenvate, Rafae Aziz, Nikhil Ravi Thiruvengadam, Yun Huang, and Aaron Steinfeld. 2011. Field trial of tiramisu: crowd-sourcing bus arrival times to spur co-design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1677–1686.