



# Envisioning and Understanding Orientations to Introspective AI: Exploring a Design Space with Meta.Aware

Nico Brand

School of Interactive Arts and  
Technology, Simon Fraser University,  
Surrey, British Columbia, Canada  
nbrand@sfu.ca

William Odom

School of Interactive Arts and  
Technology, Simon Fraser University,  
Surrey, British Columbia, Canada  
wodom@sfu.ca

Samuel Barnett

School of Interactive Arts and  
Technology, Simon Fraser University,  
Surrey, British Columbia, Canada  
sbarnett@sfu.ca

## ABSTRACT

Introspection is the practice of looking inward for ongoing self-examination. It involves considering one's past experiences and asking questions about the present and future. Our work investigates how AI could open new possibilities for supporting introspective experiences. Adopting a design fiction approach, we created a fictional company called *Meta.Aware* to contextualize 4 different Introspective AI product concepts in the form of video sketches. We used the Meta.Aware platform to conduct interviews with 17 participants, using the 4 concept videos as prompts for discussion. Participants had a range of reactions related to perceived benefits and tensions in this emerging design space. We interpret these results to outline future design directions for mobilizing AI as a resource to support introspective experiences over time, as well as to reflect on issues and dilemmas bound to this emerging design space.

## KEYWORDS

Introspection, Personal Data, AI, Design Fiction, Research through Design

## Final version published as:

Nico Brand, William Odom, and Samuel Barnett. 2023. Envisioning and Understanding Orientations to Introspective AI: Exploring a Design Space with Meta.Aware. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*, April 23–28, 2023, Hamburg, Germany. ACM, New York, NY, USA, Article 175. <https://doi.org/10.1145/3544548.3581426>  
Presented at CHI '23, April 23–28, 2023, Hamburg, Germany  
© 2023 Copyright held by the author(s).

## 1 INTRODUCTION

Introspection is the practice of mentally looking inward and examining one's own thoughts, values, and emotions [18]. Introspection is different from simply recollecting memories; it requires critically assessing key past experiences and patterns in one's life and asking questions about what has been achieved and what one wants

in the future [95]. People construct their self through an ongoing relational engagement with their things which, through their material qualities, shape how the world is perceived and understood [56, 59, 96]. In this way, people and their things are co-constituted — things play an active role in the making of the self through mediating how one contemplates the past, present, and future [5, 7, 56]. As technologies have become woven into everyday life, the kinds of things entangled with people's practices have expanded due in part to their growingly vast and diverse archives of personal data [8, 29, 37, 85]. The convergence of social, mobile, and cloud computing services have created a world in which people's life experiences are captured through explicit and implicit forms of personal data on a greater scale than ever before (e.g., [58, 78, 86]).

These shifts raise new issues for the HCI community. The emergence of vast personal data archives creates new opportunities for people to introspectively reflect on their past tastes, emotions, and experiences bound up in their digital records. Yet, the sheer scale of personal data archives also presents challenges in terms of how elements in one's life could be represented in forms that could offer valuable introspective resources.

Artificial intelligence (AI) has increasingly become of interest to design researchers (e.g., [10, 34, 64, 103]). AI offers intriguing possibilities for surfacing, exploring, and engaging with patterns across personal data archives. Yet, few examples exist in the HCI community that propose how AI might be drawn on to create new applications that support introspection. In what ways can personal data support the practice of introspection? What roles might AI play in surfacing life experiences and behavioral patterns bound up in people's data? How might a person's relation to their Introspective AI model change as they themselves continue to evolve? And what potential benefits and frictions exist in this emerging design space?

To investigate these questions, we designed a fictional company called *Meta-Aware* as well as a suite of 4 design proposals that the company offers: *Mind Probes*, *Vision Shrine*, *Hello*, *CyberSelf*, and *Dream Streams*. Each proposal extends existing introspective concepts to explore new forms of AI systems that leverage personal data records as new potential resources for introspection. For each proposal, we created a short animated video scenario [65, 68, 105] as a prototype that communicates how it works and might be used. Situating our 4 proposals in the fictional Meta.Aware company connected all concepts to a broader 'Introspective AI' model that learns from user interactions over time. Inspired by research through design fiction approaches (e.g., [4, 13, 84]), our aim is to provoke dialogue on the emerging Introspective AI design space. We conducted one-on-one interview sessions with 17 participants that had prior experience with practicing introspection. These sessions

probed on potential benefits and drawbacks of the proposals and elicited a diverse range of reactions.

Our research makes two contributions. First, it advances the HCI community's understanding of Introspective AI, alluding to future product and service forms as well as areas best avoided. This helps broaden and define the Introspective AI design space, which can be used as a generative resource for future research and practice. Second, it provides a case demonstrating how speculative design proposals can work to provide insights into current practices and inspire creative responses in the form of new design ideas. This helps support and extend HCI's techniques for leveraging design fiction as a research approach to inquire into potential technological mediations in the future.

## 2 BACKGROUND AND RELATED WORK

Related work examines the intersection of Introspection and AI as a design material.

### 2.1 Introspection and the self

Introspection is an essential part of the human condition and has been the subject of wide philosophical discussion (c.f. [18, 88, 89]). Gould's often-cited definition characterizes introspection as "an ongoing process of tracking, experiencing, and reflecting on one's own thoughts, mental images, feelings, sensations, and behaviors" [43:719]. This definition brings attention to introspection's temporal qualities — it is a continual practice across one's life that can occur in ways that are incidental and unstructured or that are deliberate. Self-introspection focuses on self-development and discovery through the process of reflecting on past and present thoughts, emotions, and experiences in relation to an, often evolving, vision of one's ideal future self [35, 79, 95].

A limited amount of research in HCI has explored how experiences of momentary contemplation, broadly construed as 'introspection', might be supported through virtual reality applications [60] or interactive art museum exhibits [87]. Our work explores the potential and limits of AI in generating digital resources to support the practice of self-introspection. We contribute insights into where different design qualities could work together to support introspective experiences over time.

### 2.2 Exploring AI as a design material

AI and Machine Learning (ML) have increasingly become accessible to HCI and design researchers. Recent work has begun to investigate how AI can be better mobilized as a material for designers (e.g., [10, 34, 41, 61, 69, 81, 102]), non-experts (e.g., [21, 104]), and various other stakeholders (e.g., [52, 53]). The 'learning' that is often core to AI refers to the iterative process of building models of phenomena in the world. While traditionally relying on static labels, training datasets, and statistical models, more recent 'human in the loop' approaches, such as interactive machine learning, focus on training models through frequent end-user interaction, where the end-user corrects or confirms an AI model's predictions [2, 104]. Interactive machine learning offers the potential to generate more accurate models [3, 42]. However, an AI model's inferences 'under the hood' can be unpredictable and unintelligible for end-users, posing key barriers to producing improved models [38, 62].

*Explainable AI* has emerged to advance strategies for making models more transparent through explanations of algorithmic decision-making processes [100]. *Explainability* has shown some early promise in setting end-user expectations by unveiling an AI model's anticipated decision-making accuracy [91, 103, 104]. Recent developments like Google's Teachable Machine project [58] enable end users to create personalized AI models from the ground up. For example, users can train their own emotional sentiments of sound samples, music, gestures, images, and so on. Although the complexity of models is limited, this project, and the broader research area of explainable AI, provoke questions about how one might 'collaborate' with machine intelligence and influence models of a user's perception of self.

Recent HCI research has also critiqued *explainability* as a design goal through situating *uncertainty* and *unpredictability* as design qualities of AI that should be explored [10, 20, 40, 54]. Yet, the uncertainty that comes with AI can make it more challenging to work with as a design material [9]. 'AI' is often a set of things (e.g., data, models, metrics, uncertainty) that, taken together, include "several forms of heterogenous materials that involve their own competencies" [20:17]. Importantly, these works argue that rather than treating qualities of uncertainty and unpredictability as problematic, they ought to be viewed as contingencies of working with AI and, in this, as creative and generative resources for design. Hsueh et al. [54] demonstrate that unpredictable, indeterminate qualities of AI can be scaffolded to enable people to interpret machine output which, in turn, can nurture more complex and meaningful interrelationships with AI applications. However, currently little is known about how such strategies might help facilitate the development of a personal introspective AI model via interactions with it over time.

Commercial products are now emerging that apply AI to support "conscious self-discovery" [1]. The Replika AI chatbot [112] produces emphatic conversations, learning from user input so they can "express and witness [themselves]." Replika leverages OpenAI's API [113] and their GPT-3 language prediction model. This autoregressive language model accepts textual prompts and generates cohesive responses, continuously adjusting its internal 'weights' — or importance scores — to inform how signals flow between nodes within its neural network structure. Since this model constantly evaluates its 'weights,' it becomes more attuned to a specific user over time, learning to create more tailored text responses. The Replika AI Chatbot encourages this training further by asking users to provide feedback if responses were relevant or not. Other products include AI-augmented journaling (e.g., [114, 115]) and dream journaling (e.g., [116]), where AI is used to organize and analyze user entries. Nearly all of these applications focus on facilitating guided introspection (i.e., similar to an audiobook) as opposed to creating new kinds of personalized resources for introspection. These applications also do not include rich accounts of personal data other than the limited input that users directly input (e.g., journal entries or images). Conversely, services such as Facebook Memories and Apple Memories use AI to curate personal data for recollecting past experiences; however, they are not created with the singular goal of prompting introspection.

There is a growing need for HCI research investigating the possibilities and perils of AI systems becoming entangled with people's

intimate everyday practices [41, 48, 64]. More research is needed that investigates alternative visions of AI applications and how they could, or should, operate [10, 28, 64, 72, 73]. Design fiction and related forms of design research are increasingly being mobilized to interrogate how emerging technologies, such as AI, shape human experience and to provide a focused context for the debate over their future potentialities (e.g., [77, 84, 92, 97]). In our work, we explore what AI might hold for creating rich introspective resources through four design proposals of fictional Introspective AI products. We aim to open critical dialogue with a diverse set of research participants on the potential benefits and dilemmas in the emerging Introspective AI design space.

### 3 METHODOLOGY

Previously, we published a pictorial that motivates the Introspective AI design space through the initial proposal of seven fictional products [16]. These design proposals appeared as 1-page visual spreads that largely focused on unpacking what each concept ‘is’, which was paired with a brief use case scenario. We extend this work in three important ways. First, we substantially developed four design proposals – *Mind Probes*, *Vision Shrine*, *Hello Cyberself*, and *Dream Streams* – by designing animated videos, developing a detailed scenario script, and working with voice actors. This enabled us to unpack each concept in a far more detailed and resolved format. We selected these four design proposals as they form a diverse set, where each proposal builds on a different preexisting introspective concept or method and extends them through applying AI in a novel way. Additionally, in our view, the other three remaining proposals originally detailed in our prior research were either situated too close to already existing products – thus, limiting the capacity for speculation and debate – or they were too far-fetched and complex to encapsulate within a relatively brief, intelligible video scenario.

Second, we developed a fictional company, *Meta.Aware*, to contextualize these Introspective AI design proposals within a coherent world narrative. Third, we used the *Meta.Aware* company video with our set of 4 design proposal videos in an elicitation study with 17 participants. We conducted one-on-one interview sessions that probed participant reactions and prompted debate on the benefits and drawbacks of each. Next, we summarize key parts of our design process to highlight important qualities of our Introspective AI proposal videos and then describe participants, data collection, and analysis.

#### 3.1 Meta.Aware: Creating Context for an Alternative Present

We designed the fictional company *Meta.Aware* to contextualize the origin and existence for all 4 design proposals. This enabled us to connect the design proposals with a similar design language, and it emphasized that the ‘Introspective AI’ model learns from user interactions in and across each of the product concepts. This decision builds on a trajectory of HCI and design research that has mobilized design fiction as an approach to speculate and provoke dialogue on the social, cultural, political, and ethical implications of future technologies (e.g., [4, 12–14, 26, 67]). Following Coulton et al.’s characterization of design fiction as “. . . collections of artifacts that, when viewed together, build a fictional world” [26:4], our decision to frame the design proposals in the context of *Meta.Aware* was important for collectively integrating them within a cohesive narrative, potentially helping viewers suspend disbelief.

*Meta.Aware* is introduced through a video that portrays an interview with its founder and CEO, where they frame the company as “harnessing the power of artificial intelligence to leverage your personal data to empower you to develop your own practices and rituals for introspection.” We designed the CEO’s script to

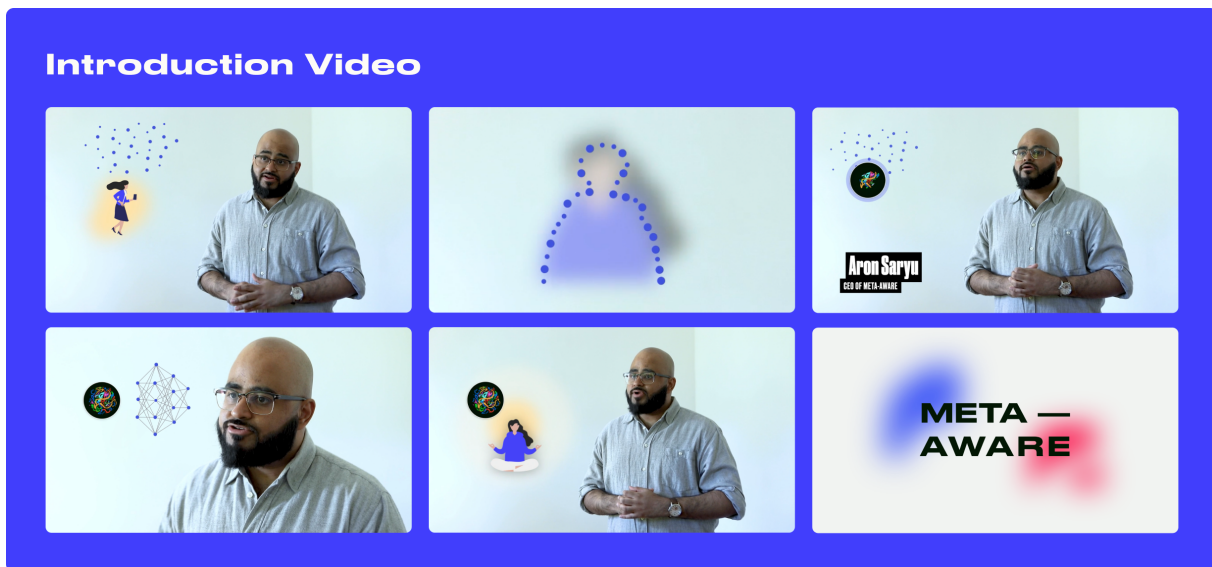


Figure 1: Images from the introductory video where *Meta.Aware* is introduced as a company offering the suite of Introspective AI services explored in our study, and Alison is presented as the main protagonist in the video scenarios.

balance a sense of optimism and wariness about current technology (see Figure 1 and supplementary video figure). The narrative tone of the CEO’s script aims to probe on the paradoxical relationship between extractive technologies that may exploit user data (c.f., [27, 28, 106]) and the potential for beneficial introspective practices to be supported. We want to explore perceived tradeoffs, benefits, and tensions in the emerging Introspective AI design space. The Meta.Aware video also alludes to the two main characters that appear across all videos: *Alison*, a mid-20s professional living in North America, and *Alison’s Introspective AI model*, which drives the four design concept proposals, continually learning from data collected from Alison’s life to generate introspective resources.

Through our framing of Meta.Aware, we wanted to move away from portraying AI as a single ‘all-knowing’ agent that takes on human-like reasoning or a human-like form. Similar to recent work (e.g., [10, 20, 41, 54]), we see the uncertainty that comes with this difference as a generative opportunity for design practice. We situated Meta.Aware’s AI as a context-aware agent that mediates interactions between Alison and her personal data by learning and making inferences about her life through situated introspective prompts. We also wanted to explore ‘mistakes’ that an Introspective AI might make and the different ways that these imperfections or misjudgments in Alison’s model might be perceived.

### 3.2 Design Proposals: A Suite of Introspective AI Products

For each of our four concepts, we created a short video prototype. Each video combines a mix of narrated descriptions and user scenarios based around Alison interacting with her Introspective AI through a specific Meta.Aware product. Our approach builds HCI research that uses video sketches to express “ideas and use situations that would be impractical or impossible to create in conventional prototyping techniques” [65:23] when investigating the potential implications of emerging technologies [68, 70, 98, 101, 105]. We tailored the scenarios around key questions and topics we were interested in exploring with participants. Our approach shares similarities with prior HCI research that has argued for exploring the potential roles, values, and social boundaries of emerging technologies by using multiple design proposals to establish a better overall understanding of the design space [23, 32, 36, 46, 77, 97]. While diverse, these works show the value of encouraging participants to imagine future interactions and react to them by drawing on their own experiences. Next, we detail the four design proposals and the underlying issues each aim to explore. We encourage readers to view the design proposal videos attached as supplementary materials as a part of this submission.

1. *Mind Probes* is a smartphone app that works in tandem with external hardware sensors: sound, color, smell, haptic, and vision

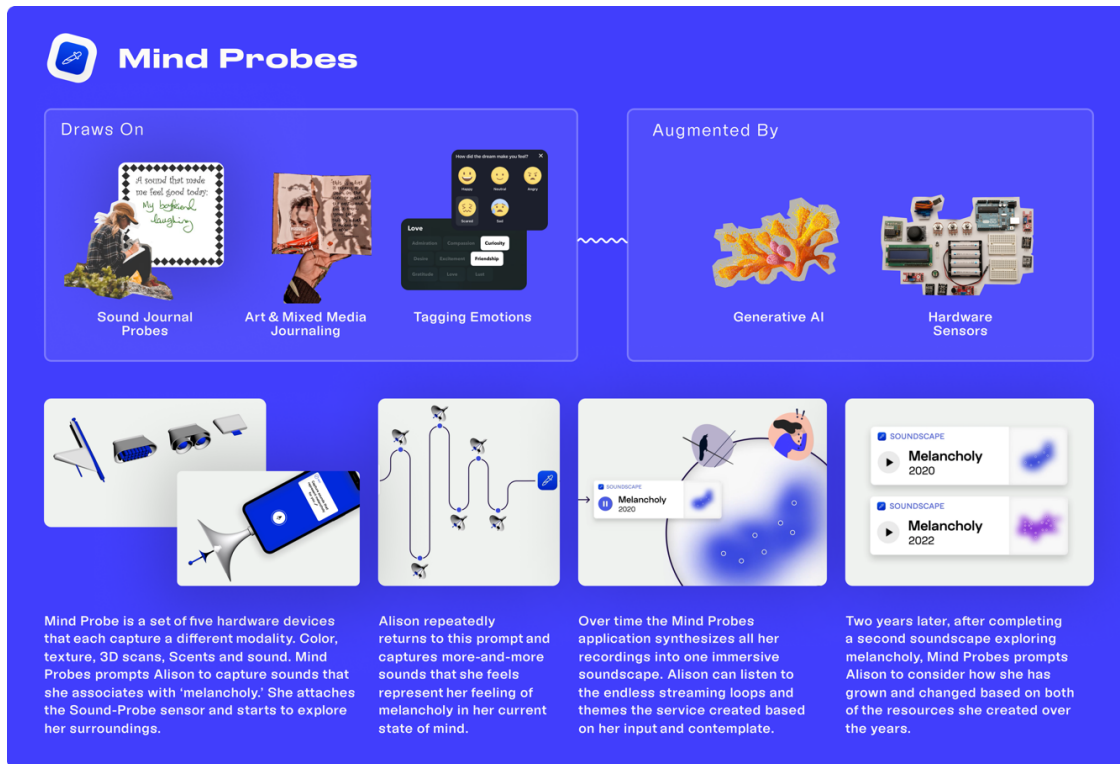


Figure 2: Top Row, Left to Right: Mind Probes draws inspiration from the introspective practices of sound, art, and mixed media journaling, as well as tagging emotional sentiments for journal entries; it explores how generative AI and hardware sensors might extend these practices. Bottom Row: Abbreviated storyboard of the Mind Probes scenario.

(see Figure 2 and supplementary video figure). It encourages introspection through activities akin to a scavenger hunt. Mind Probes prompts users with questions linked to specific sensory modalities, asking them to collect sensory stimuli from the material world that reflect social and emotional associations — connecting inward associations with encountered phenomena. In our video scenario, Alison uses the sound attachment to collect recordings related to her perception of ‘melancholy.’ The system generates an immersive soundscape based on Alison’s input. Over time when multiple soundscapes are created, the service offers to compare these resources to contemplate how she and her emotions changed.

This proposal is inspired by creative prompts about one’s perceived environments (e.g., sounds, colors, objects, etc.) from the influential self-help book “The Artists Way” [117], art and mixed media journaling [118], and tagging emotions over time on digital journaling applications [114, 119]. We saw potential in drawing on artistic approaches to journaling as they focus on emotive and expressive ways to contemplate one’s mental state. In addition, we wanted to explore alternative avenues for people to track and tag emotions over time as current approaches appear one-dimensional and limited (e.g., most current commercial applications ask the end-user to select one of five emotions to add a sentiment to a journal entry). Finally, we also explore novel hardware sensors that,

over time, enable the capture and creation of unique, multi-sensory journal entries. In this way, Mind Probes draws on AI’s generative potential to collaboratively create representations from data that users are generating. Here, we draw inspiration from technology that lets users train their personal models [21, 120, 121] as well as generative soundscapes informed by personal data [15, 122, 123].

As a person encounters new experiences across different stages of life, elements of their identity may stabilize while others could transform [7, 88, 95]. However, the subtle, often unpredictable qualities of personal growth could make it challenging for the AI to notice and adapt to, thus potentially requiring the user to explain them through data-producing activities. This proposal explores questions including:

- How would an Introspective AI model that is taught with rich and subjective forms of personal data be perceived?
- How would people react to a non-human agent that helps track specific emotions over time?
- How willing would people be to perform the labor of ‘collecting’ data for their Introspective AI?

2. **Vision Shrine** is a tangible device that visually manifests a user’s goals, dreams, and desires as data collages—an ‘ideal self-canvas’ that updates in real-time (see Figure 3 and supplementary



Figure 3: Top Row, Left to Right: Vision Shrine draws inspiration from Vision Boarding and Self-Discrepancy Theory and augments these introspective methods with a curatorial AI agent that leverages Alison’s personal data to track and train the model. Bottom Row: Abbreviated storyboard of the Vision Shrine scenario.

video figure). Vision Shrine adopts a ludic framing [39] through playful confrontations that change a user's priorities on their canvas based on their behavior, raising questions around the ranking of personal desires in the context of one's perceived ideal self.

The Vision Shrine design is inspired by two introspective practices. First, *vision boarding*, an activity where one visualizes desires as collages on a pinboard. Here, people consider what they want to achieve in life and create a visual manifestation that helps them stay on track with this vision [17, 107]. Vision shrine augments this analog practice which can feel tedious and static. Second, *Self-Discrepancy theory*, a taxonomy that helps to unveil internal tensions among one's ideal, ought, and actual sense of self [51, 89]. Vision Shrine extends these methods into a technologically mediated introspective practice, helping users to dynamically examine and visualize those discrepancies to understand internal conflicts better and dive deeper into their roots.

Vision Shrine probes into how a *curatorial algorithm* [46, 109] might enact forms of agency as it playfully alters one's visions on the canvas. This AI-enforced 'reality check' is intended to explore where people's boundaries lie if technology actively plays a role in their rituals of self-development. It also explores playfulness as a strategy to support ways to enable end users to *teach* the

machine through implicit interactions. For example, interactions like re-sizing or re-positioning of visions are ways to train the underlying algorithms, thus subtly reinforcing images users feel more represented by. This proposal explores questions including:

- How acceptable might applications that critically challenge individuals' current desires in relation to their various future goals be? Where could perceived frictions or benefits exist?
- To what extent could a curatorial system like Vision Shrine lead to perceived (in)authentic visions of one's future?

3. *Hello, Cyberself* offers a conversational window into the assumptions (and biases) that a personal Introspective AI model has developed over time. It leverages real-time voice cloning technology [25, 57] to speak to you in your own voice (see Figure 4 and supplementary video figure), expressing prompts to *you as you* – embodying your personality traits and then revealing the data 'under the hood' that generated these inferences. Hello, Cyberself is inspired by personality tests that leverage the 16 personality types of the Myers-Briggs Type Indicator [124] or the big five personality traits [74]. Here, people commonly fill out an extensive set of questions about how they perceive themselves in a given scenario and receive a detailed analysis of the assigned personality type. Today, several commercially used algorithms and APIs promise to extract



Figure 4: Top Row, From Left to Right: Hello, Cyberself draws inspiration from personality tests and confronting cognitive bias, and explores how they could be extended through confronting 'one's self' and, subsequently, understanding (and potentially editing) inferences that the model relied on to generate each Cyberself report. Bottom Row: Abbreviated storyboard of Hello, Cyberself scenario.

personality insights about users in order to gain information about customer behavior and optimize targeted advertisement [125, 126].

Critical of this largely being hidden from users, we were inspired to create a service that offers a way to make similar assertions visible for introspection and self-insight. The Hello, Cyberself concept is also derived from actively examining and exposing cognitive biases [80] and looking into the opportunity of AI as an observer that presents confrontational perspectives prompting us to investigate our own perspectives more critically. Hello, Cyberself enables users to confront factors that shape how their model formulates assumptions that fuel its behavior and tweak them if desired, thus, exploring methods of machine interpretability [11, 38] and teachable machine learning [21, 120, 121]. This proposal explores questions including:

- What benefits and tensions might emerge from encountering a digitally-mediated imprint of one’s self?
- How might people perceive their identity and sense of self once they are challenged by a machine?
- How would people react to having an added sense of transparency and the ability to ‘edit’ or ‘correct’ their personal model?

**4. Dream Streams** combines a tangible device paired with mobile applications to offer windows into one’s subconscious and open new pathways to self-awareness (see Figure 5 and supplementary video figure). Dreams offer a way for your subconscious mind to communicate with your conscious self and offer a significant window into phenomena that shape our innermost desires, fears, and goals [31, 83]. As such, dreams can offer important resources for introspection [60, 99]. This proposal speculates on how an Introspective AI could mobilize sleep and active recall of dream experiences to generate introspective resources. The device assimilates data from sleep-tracking devices with dream journal recordings and creates three distinct introspective experiences: An abstract audio-visual experience, pattern analysis, and a Spotify playlist.

With this proposal, our goal is to probe on how to engage with sleep data in ways that extend beyond common forms of quantification (e.g., dashboards and graphs) and might support novel interactions that allow for ambiguous, ethereal, and interpretive experiences with people’s sleep data archives. We aimed to envision a holistic system that revolves around the intentionality of the device’s situation within a user’s bedroom. The system is designed to foster the daily ritual of dream journaling, minimizing distractions that people are usually exposed to (e.g., compared to

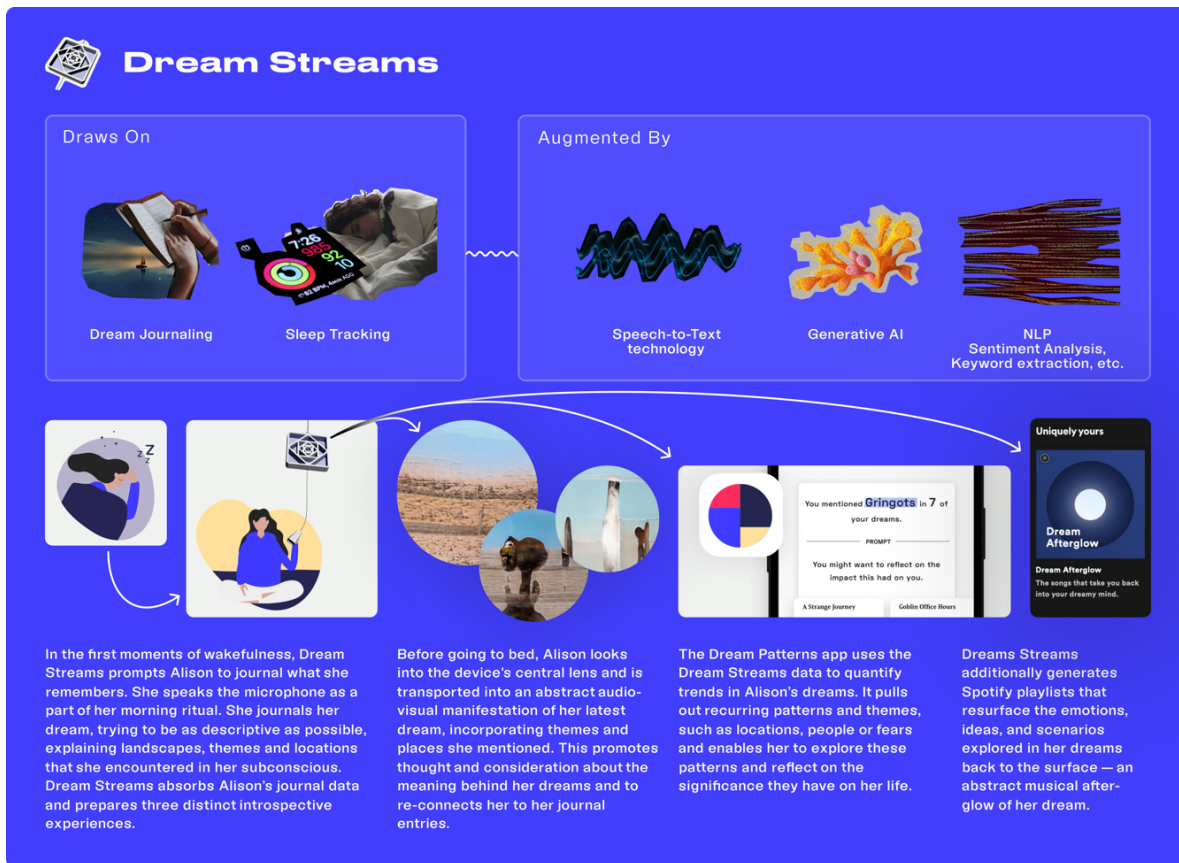


Figure 5: Top Row, Left to Right: Dream Streams explores the intersection of dream journaling and sleep tracking, augmenting them through several AI technologies to generate a suite of Introspective AI resources. Bottom Row: Abbreviated storyboard of Dream Streams scenario.

journaling on a smartphone app). Further, we wanted to explore the potential of generative AI in helping people to express and remember their dreams based on their unique descriptions – a notion that is more accessible by the day with the current explosion of commercially available AI tools and APIs for image and video [127–129]. Dream Streams also use speech-to-text technology [111] and natural language processing [108], leveraging features like sentiment analysis, keyword extraction, as well as classifying names and places. This proposal explores questions including:

- What might be revealed through looking inward at one’s self at sleep?
- Could unknown or forgotten personality traits and orientations be surfaced?
- How beneficial would people perceive an AI’s interpretation of their dreams as an introspective resource?

### 3.3 Elicitation Study, Participants, and Analysis

We used the Meta.Aware introductory video and our set of four design proposals in an elicitation study with 17 participants. We showed participants the videos during one-on-one semi-structured interview sessions that also touched on their own introspective practices. All interview sessions were conducted through a secure video conference system.

**3.3.1 Recruitment and participants.** We recruited a diverse sample of people to elicit a wide range of discussions about and beyond our design proposals. A total of 17 participants were recruited through word of mouth and advertisements from a large city in Western Canada. 12 participants self-identified as female, 4 as male, and 1 as non-binary. The median age of our participant pool was 27 and with an average age of 30; the youngest participant was in their early-20s and the oldest in their mid-60s.

Occupations included Filmmaker, Artist, Big Data Analyst, Spiritual Counselor, Environmental Education, Professor of Human Geography, UX designer, and AI & UX Researcher. Regarding their practice of introspection, 11 were frequent practitioners, 4 moderate, and 2 communicated interest but did not practice often. This was the general self-reported approach to introspection, but during the interviews, participants elaborated on their respective approaches, displaying, at times, quite nuanced practices. Our diverse sample is not meant to be statistically representative but instead generative.

**3.3.2 Procedure.** The study sessions consisted of 1) an introductory interview to understand each participant’s background; 2) an introduction to the fictional company Meta.Aware; 3) viewing each concept in sequential order 1-4; and 4) a final longer-form interview. In the introductory interview, we asked participants to discuss their introspection practice and its significance in their life. We asked them to reflect on their use of technology, what kinds of personal data they generate and possess, and what the term “AI” means to them. Then, we showed the Meta.Aware introductory video. For each design proposal, we showed the video via a YouTube playlist we created. After each video, we asked if the concept and scenario were clearly understood and then probed participants’ first impressions, asking them what they felt about different aspects of the use scenario.

Following Odom et al. [77] and Vitale et al.’s approach [98], in the final interview, we asked participants to reflect on and across their experiences of all the design proposals. We asked them to pick the most or least valuable for them, discuss the most positive or negative aspects across all the concepts, and elaborate on the ideas behind them based on how they would fit their needs and values. One member of the research team led the interview, while 1-2 additional members observed, took notes, and posed follow-up questions. All interviews were conducted in English via a video conferencing system. Interviews lasted between 120 – 180 minutes



Figure 6: A snapshot of research participants extracted from the video calls.



and were video recorded. Participants received \$50 as compensation for their time. All participants visible in Figure 6 consented to have their images appear in research publications.

**3.3.3 Data analysis.** To analyze the data, we used Braun and Clarke’s approach to thematic analysis [24]. Our interviews produced a total of 36+ hours of video recordings. Recordings were first transcribed using the Otter.ai automatic transcription service. We then conducted a subsequent round of reviewing all transcriptions to edit or correct errors (e.g., typos) inadvertently generated by the service. Throughout the data analysis process, two members of the research team operated as coders and used a hybrid approach that first involved deductive and then inductive coding. The research team coded the data and discussed the themes and interpretations with other authors during weekly meetings over the course of four months. We grouped codes into categories and identified themes across categories. Our process was iterative and included searching for themes that led to shifting thematic patterns across our data that eventually stabilized [71]. As primary themes in the data stabilized, it became progressively clear how and where relational subthemes could be nested within the primary themes. We also created affinity diagrams to model and reveal connections and differences among participants and across design proposals.

## 4 FINDINGS

Next, we present the results of our analysis which are distilled into three primary themes. In the first two themes of analysis, we contrast opinions on what role technology should have in supporting introspection: numerous participants recognized the value in engaging in acts of dynamic co-creation with Introspective AI (IAI) services while also reflecting on their limits and sharing creative control (Theme 1). Further, frictions emerged around the IAI overstepping boundaries of acceptability when it took on more direct, at times brash, roles to confront users with new perspectives on their life (Theme 2). Then, we synthesize findings that illustrate the longer-term place, scalability, and potential for new rituals with IAI envisioned by participants (Theme 3).

### 4.1 Theme 1: Co-Creation with new introspective resources

The first theme captures participants’ reactions toward the co-creative qualities of the four design proposals. Participants perceived co-creation as an ongoing practice of contributing personal data to the IAI in exchange for new introspective resources. As our study progressed, it became apparent that participants also recognized how the process of reviewing and manipulating data—which teaches and augments the IAI— can also be seen as form of co-creation. Taken together, the following reactions highlight the perceived value of Introspective AI in an active role as a co-creator, as well as tensions around the limits of this “creative agency” in a user’s journey of self-exploration.

**4.1.1 Contributing data and anticipating results.** *Mind Probes* prompts users to contribute data probes such as sound snippets, colors, textures, and scents as a response to prompts (e.g., “Capture sounds that represent melancholy to you”) and subsequently generates interpretive resources that manifest the user’s responses.

Aligned with many other participants, **P14**, a feminist filmmaker and textile artist who likes to be generative in their introspective practice, valued the notion of creating resources with the IAI and speaking emotions “*into existence . . . [rather than] just sitting there and ruminating with your anxious brain.*” Some participants valued the act of collecting data probes as an introspective practice in and of itself. **P2**, a UX Strategist with a long-term analog journaling practice, noticed the potential for *Mind Probes* to educate people and help them express their needs: “*I think it’s training someone to be more emotionally intelligent . . . Identifying with emotions and associating them with something like sounds, scents, objects is powerful.*”

Several participants elaborated on the creative activity of responding to a *Mind Probes* prompt as an opportunity to create new narratives based on whom they are at the moment, helping them to move toward whom they want to be in the future. Participants also saw the act of data collection as generating anticipation for an introspective resource that would eventually be produced. **P12**, a filmmaker who defines introspection as: “becoming aware of the stories we tell ourselves on a daily basis,” was excited about this quality of *Mind Probes*: “*To me, the feedback is at the end of those combinations of sounds and that you’ve created something. It’s kind of a scrapbook, in a way, but one that’s just of your senses and where you can look back to over the years. That would be the reward for me.*”

As exemplified in **P12**’s passage, many participants perceived the act of collecting data as enabling them to play a more active role in co-creating resources with the IAI. The sentiment that creative agency could be positively reinforced through having first-hand control over which data would be added to the IAI model surfaced as an important theme, which often emerged across discussions of the design proposals. **P12** captures this sentiment well, appreciating *Vision Shrine*, *Mind Probes*, and *Dream Streams* as services that made them feel like a “co-creator in the experience” and “asking for a creative process,” which enhanced their perceived level of comfort with using such AI-augmented services.

**4.1.2 Balancing curatorial control.** *Vision Shrine* prompts users to manifest current emotions in the form of spoken journal entries that capture their desires, dreams, and life ambitions and populates a digital vision board based on their input. Participants generally appreciated that the *Vision Shrine* helps to give form to internal processes in tangible ways – a theme summed up well by **P1**: “*You’re just putting it out into the world and saying it and then it’s on your board. I think that really changes how you think about things, and how scary they seem, or actually being motivated to go for it*” (**P1**).

In contrast to *Mind Probes*, *Vision Shrine* plays a more active role in co-creating introspective resources by meddling with the content on the display. The service rearranges and changes the size of the displayed “visions” in response to a live stream of the user’s data (e.g., cataloging what they actually spend time on) to create an ongoing visual dialogue between a user’s lived reality and their envisioned ‘ideal self.’ We anticipated strong negative reactions towards a service that attempts to influence a user’s visions directly; however, instead, we uncovered an interesting tension. Several participants both appreciated *Vision Shrine*’s confrontational and ludic qualities while also expressing discomfort with it. **P7**, a

professor in Human geography, exemplifies this dichotomy: “What you’re really demonstrating here is visualizing cognitive dissonance, which I think is something that we’re all suffering from. Basically, it [Vision Shrine] says, well, you’re saying one thing, you’re wanting one thing, and you’re doing the other thing. It’s just making that really aware. So it’s a cognitive dissonance Big Brother.” For **P14** Vision Shrine is “kind of passive aggressive, but in a funny way” and acknowledged that the way it asserts its curatorial agency needs to be nuanced: “it’s the balance between being self-disciplined and feeling like you’re being parented.”

This perceived tradeoff was common in our interviews, often leading to interesting discussions about the value of a service that holds users accountable in playful ways and how such a confrontational approach needs to navigate the tensions of feeling productive while not undermining a user’s introspective practice. **P1** described how they desired to preserve some key elements on their *Vision Shrine*, drawing on the tension that sometimes they might still think certain elements are important even if not reflected in their daily behavior. This dialogue highlights the need to balance user desire for more creative control without losing the *Vision Shrine*’s focus on holding one accountable to their self-discrepancies biases [89]. **P1** proposed new interactions that would enable a user to “pin” visions that they did not want to see fading away or to be able to set a parameter for the size that the representation of a given goal could not shrink beyond. They speculated that an alternative version of *Vision Shrine* might not adjust the scale of a user’s goals at all, instead drawing attention to them more passively through wiggles or pulses to prompt a user to reflect on and engage with their various competing life goals. This desire was echoed by participants who felt apprehensive about *Vision Shrine*’s curatorial influence and expressed a desire to have total creative control while keeping the “reminders” to stay on track. This sentiment is exemplified by **P10**, an invested practitioner who likes to create resources that inspire her and help her with decisions: “these are my goals, and then they are evolving and that’s being inputted, but I would still want to be playing an active role in that. I think that it wouldn’t be something I would like the AI to be involved in, besides reminding me to stay on track.”

**4.1.3 Manipulation and revision as co-creation.** All four services in our design proposals are imbued with interactions that are intended to train the IAI model to make ‘better’ decisions when generating new introspective resources. Since a user’s envisioned ideal self can change over time, we wanted to interrogate how their IAI model might adjust to such changes. Interestingly, to our surprise, participants generally viewed interactions where a user and IAI service work to add, subtract, or otherwise manipulate their digital algorithmic profile as creative acts that could nurture the growing relationship between a user and their AI. **P7** valued how *Hello, Cyberself* offered a “view under the hood” that enabled a user to ‘correct’ its data points: “Anytime the agency’s back on the person to work with the data, [...] that ability to disturb or to reprioritize, or to rethink, and change your mind, those are moments that I like, within the context of building relations between your AI and yourself.”

Yet, some participants also initiated concerns over this act and questioned if correcting the model could lead to feelings of inauthenticity: “If you change the data point, then it’s not actually

accurate. You’re just cheating yourself. So it’s up to the user, the only person that they’re cheating is themselves.” (**P12**). Interestingly, **P1** desired to creatively manipulate the model’s parameters to see how certain results would change. For example, they speculated on the idea of a slider that lets a user adjust the level of “boldness” that assertions are presented; they desired to experience a sequence of prompts based on one adjustment and then set the slider to an extreme and re-visit the whole sequence, hopefully picking up new insights through this process.

Aligned with other participants, these findings show a desire to tweak and exaggerate parameters to better understand who they are in relation to the more magnified views that the IAI model may show. **P1**’s passage highlights an unexpected interpretation of how revising and training a model can operate as a form of co-creation and a form of introspective self-exploration.

## 4.2 Theme 2: New perspectives and confrontations of self

The second theme reports on participants’ reactions to the qualities of our design proposals that introduced a user to new perspectives through personal data in alternative, often challenging, ways. It foregrounds the fine line of IAI services needing to balance productive confrontations and new perspectives with overstepping personal boundaries.

**4.2.1 Distrusting analytic perspectives.** Participants voiced concerns about services that seemed to analyze and ‘pre-synthesize’ their data too heavily. **P5**, an AI researcher concerned with the effects of AI biases, expressed her reluctance towards how *Hello, Cyberself* generates targeted assertions based on data: “Patterns are dangerous, because they are narratives, they are not unbiased or unvarnished fact.” Yet, aligned with many other participants, **P5** also acknowledged the potential benefit of an IAI that is able to locate and curate novel connections in an introspective practitioner’s data to gain new perspectives: “I like the idea of helping with pattern recognition for things that we don’t recognize in the moment, such that we can reflect upon them in moments of privacy or respite later. . . . Things that you normally wouldn’t pay attention to.” **P5**’s reflection represents a consistent theme—participants were apprehensive of data being analyzed yet appreciated the promise of an IAI generating new introspective resources.

Over the course of our study, it became clear that participants felt most concerned when they perceived an IAI service appeared to interpret, analyze, or act on these surfaced data patterns instead of simply displaying and drawing attention to them. These reactions often pointed to the importance of the presentation of patterns in data such that users are not pressured to feel like they are competing with assertions the system made: “I don’t want to feel like my data knows better than me” (**P12**). Rather, participants desired services that offered a space for them to interpret patterns and derive meaning on their own terms after new ideas or new vantage points had been introduced. A strategy that was perceived to alleviate this tension emerged in discussions on how *Hello, Cyberself* offered a view ‘under the hood’ of the IAI by presenting explanations of how assertions and patterns were created, which **P5** valued and wanted to see further developed: “The idea of showing [the data] could be integrated more into the forefront; essentially, showing how the sausage

is made upfront, like, we're noticing these patterns, we're not going to ascribe meaning to them immediately and leave the synthesis to the human."

**4.2.2 Confrontational Perspectives.** Inward confrontation of one's self and potential biases is key to introspection [90]. Our design proposals presented personal data and assumptions in the form of confrontational prompts that were aimed to provoke reflection. This quality appeared most lively in *Hello, Cyberself* by creating confrontational assertions about the user through a cloned imitation of their own voice. This intimate confrontation aims to prompt a spoken response from the user as they reflect on each statement—an animated dialogue with your 'digital self.' We anticipated strong negative reactions towards how *Hello, Cyberself* confrontationally invades the user's personal space. Yet, surprisingly participants generally liked the brashness of the service. **P11**, a filmmaker who uses AI as part of his artistic practice, appreciated the bold narrative tone, describing it as more valuable self-dialog compared to how "low impact" affirmative statements were on examining and challenging his personal beliefs. **P12** further hints at the opportunity of *Hello, Cyberself* in introducing confrontational statements that come from a non-human perspective: "It's interesting because it feels like it might be easier to not take it personally, because it's not a real person . . . because it's not a person that you can have your feelings hurt by."

Beyond the opportunity of such non-human perspectives, other participants were drawn to how *Hello, Cyberself* offers an unbiased perspective on a user's behaviors by merely providing a window into their personal data. **P10** contrasted this approach with conventional journaling: "You can't hide from the behaviors that you've been exhibiting. *Hello, Cyberself* shows me clearly that I'm repeatedly doing something I don't like, whereas in my journal, I can get away with it. There's no interaction, there's nothing being reflected back."

**4.2.3 Taking perspectives in—or not.** Through our study, it became clear that some participants felt uneasy about technology introducing bold claims about their personality based on data. Interestingly, most of the participants that voiced this concern did so in favor of protecting people that might be less experienced introspective practitioners. **P9**, who works in education and has a daily journaling practice, hinted at the potential for such statements to become "self-fulfilling prophecies" for people starting their journey of self-discovery. Yet, for our more advanced participants that already used technology in their daily rituals, this issue seemed less troublesome. For example, **P13**, an advanced anthroposophical practitioner and educator, had an interesting stance on how strong she would 'consider' *Hello, Cyberself* provocations: "I wouldn't mistake it for a reflection of who I am. I think that's where the distinction lies. I would think of it exclusively [as] this is the '[P13]-self' that has been created based on the data that I have been generating."

In addition to this "take-it-or-leave-it" approach expressed in previous statements, **P17**, an experienced spiritual councilor of over thirty years, started to speculate on her desire to integrate appropriate guidance on how to use resources like [*Hello, Cyberself*] more soundly—a strategy they felt was generally missing in many digital self-help apps today: "I think with this app [*Hello, Cyberself*], you should put some sort of waiver or disclaimer saying: Don't make this a false god or an idol, do not give away your own authority to this.

*This is just a tool, and you can bless it to be in service, and the only person it's in service to is you."*

**4.2.4 Abstract & Interpretable Perspectives.** Design proposals that produced resources that leveraged data in abstract ways were generally well received across participant responses: "I like the visual-auditory, abstract representations of a lot of these ideas because they're thought-provoking without feeling too inquisitive or accusatory. It's more like checking what *Meta.Aware* made from your thoughts, and then you get to decide how much of it you buy into." (**P5**). Of these services, *Dream Streams* created the most abstract resources based on a user's *dream data* (e.g., AI-generated videos based on journal entries and sleep-tracking data). Here, **P5** appreciated the inspirational and interpretive qualities of these forms: "The notion of getting a weird, hard-to-interpret output from a weird and hard to interpret and ephemeral process, like dreaming, is a very fun way of remembering to be thoughtful." More profoundly, participants voiced their desire to move beyond ways of rational self-analysis as they seemed to be unsatisfied with their current reflective practices with personal data (e.g., via journaling and self-tracking apps). **P11**, an artist and filmmaker, uses his smartwatch's quantified sleep and exercise data in his introspective practice. Although he finds value in these data points, he appreciated the surreal nature of the artifacts generated by *Mind Probes*, and expressed a desire for more abstract and playful representations of personal data in e-health services "there's not enough focus on more quirky, mystical, surreal ways to access to unconsciousness."

### 4.3 Theme 3: Desire for Presence, Scalability, and new Rituals over time

In themes 1 and 2, we have discussed a range of reactions to the concepts that encompass positive and controversial perceptions. The final theme details how participants engaged with the design proposals to envision the longer-term place, scalability, and potential for new rituals with IAI applications. Often these discussions emerged toward the end of interview sessions after participants had been exposed to all four design proposals.

**4.3.1 Supporting Presence and Habits.** Participants strongly appreciated that *Dream Streams* and *Vision Shrine* were realized as physical objects with a specific purpose and speculated on their presence in the home. **P2**, who uses the *Headspace* app as a part of their introspective practice, explained fatigue emerging from relying on their phone and a need for intentionality: "I hate feeling stuck to my phone and always having to use my phone. But what's beautiful about this is it's just there in your home, and it's there for you and becomes a centerpiece." Participants also connected the tangible quality of the IAI products as being rich mediators of introspective rituals. **P13**, who has been keeping a dream journal for 6 years, emphasized how important it is for her to have a physical object that helps to continuously facilitate a ritual through intentionality and aesthetics: "I feel like the aesthetics and the beauty of something, creates a kind of mood, it creates an emotional connection to it, that already transports you into that mindset of like, 'okay, now I'm going to do something very spiritual.'" **P13's** reflection hints at the challenge of finding an appropriate design language for devices like

Dream Streams that combine highly technical aspects with ideas and themes that are more mystical and metaphysical.

**4.3.2 Supporting Temporal Connections and Trajectories.** All IAI services incorporate different strategies to surface temporal connections across a user's data. *Vision Shrine* applies temporality on a very visible level as it responds in real-time to the input of a user's personal data to actively manifest a changing image of their dreams and desires. Throughout the study, participants recognized the value of these temporal qualities: *"I think it's interesting how it will be able to track what you've said and thought about yourself two months ago to now because we're constantly changing our narrative and our stories"* (P12). This points to participants perceiving resources like *Vision Shrine* as a potential historical archive of personal data where their recorded goals were seen as valuable introspective resources. To this end, P11 speculated on how *Vision Shrine* might help to *"loosen your fixed idea of self"* and reveal how *"past obsessions"* might no longer be relevant.

While *Vision Shrine* subtly yet continually changes dynamically, *Mind Probes* enables users to create sensory snapshots of their lived present. P3, who has an extensive daily meditation routine, acknowledged the years of time she devoted to her vipassana meditation practice and felt that the benefit from *Mind Probes* would emerge across a longer time scale: *"I feel like something like that would have to take place over a very long duration of time, but I think it would be interesting . . . how your self-narration might differ from one year to the next, or what your life story might be."*

P17 speculated on how *Vision Shrine* could be extended to show patterns across temporal ranges to support the development of self-knowledge. More specifically, she was interested in how an IAI might re-surface journal entries over time, connecting this to successful techniques she uses with clients in their counseling practice: *"People are always mystified at either the wisdom that came through them, or a different perspective that they hadn't thought of, but they need the distance to really see it with new fresh eyes."* P17 continued to speculate on an AI algorithm that could be trained to detect beautiful sound bites or deep introspective thoughts that a user expressed that would then prompt them again after time had passed. She speculated that a service like this would help people to recognize their personal greatness and wisdom. Thus, P17 hints at a less tech-intensive algorithm that wouldn't create 'judgment' but rather: *"An AI, saying 'here, this is something you yourself created. This is you, yourself. Let me help you remember these truths."*

**4.3.3 Sharing Introspective Resources.** Although the design proposals focus on a sole person's introspection with an IAI, participants identified multiple avenues for the IAI to also mediate human-to-human interaction. They speculated that a preferred interaction with the IAI would be to connect them with another person (be it a family member, friend, or medical health professional), as opposed to only providing its own analysis of the data and suggesting a private course of action that could be undertaken. Some participants felt that it would be easier to develop habits of use around a service that requires reoccurring interaction (such as *Mind Probes*) if they were to complete them in tandem with family members or trusted friends.

Beyond social accountability, many participants felt that the introspective resources generated by *Mind Probes* (e.g., synthesized

soundscapes), *Vision Shrine* (e.g., the current vision board canvas), and *Hello, Cyberself* (e.g., confrontational or amusing assertions) could also be beneficial to share and compare with others. Participants valued how these actions might reveal new commonalities and differences that could contextualize their lived experience: *"These technologies provide more of an understanding of how I exist socially, and I would still be interested in that, or how I might exist contextually to others"* (P3). While a social positioning might contribute to the development of self-knowledge, so too might a socio-historical positioning. P17 reflected on the 'heirloom potential' of the *Mind Probes* resources, imagining if she had access to their grandfather's IAI model: *"when somebody has passed, and they want to leave an archive, to their children or grandchildren, you could see what made grandpa melancholy."*

**4.3.4 More than just a user.** The IAI services are predicated on the collection and use of highly sensitive personal data. We framed the services as belonging to an opportunistic startup, as a design fiction strategy to develop immersion and worldbuilding. Some participants voiced their concern about how data-intensive IAI services might fuel further exploitation and inequities within the current data economy and, hence, desired a more positive future. P13 expressed her frustration with the state of current algorithms and the underlying commodifying infrastructure: *"I feel like there's such a strong kind of market logic in the ways in which health and wellness has been commercialized [ . . . ] The algorithms that have been working with capitalist logic are unsatisfying or will, at some point, be underperforming to our needs. Our needs are an altruistic community. We need some feeling of connection with other people that is genuine and authentic."*

This sentiment extended into considering how some participants perceived aspects of the *Meta-Aware* resources, worried that they may be yet another way for the tech industry to hijack attention. Participants also voiced their desires to participate as more active stakeholders in the development of future IAI services that they may use. P11, interested in decentralized infrastructures and WEB 3.0, expressed his wish for IAI services to be built around a community of users and developers that are in dialogue with each other because the *Meta.Aware* products *"feel like a laboratory for research into a user's mind."* Yet, he addresses this privacy concern with an idea of how this could be inverted: *"I don't want to feel like just the user. I want to feel like a researcher. That's why I think being a stakeholder in the app would help me."* P11's statement alludes to a profound need to be more than merely a user of a potential future IAI service but also actively involved in determining its ethical compass and goals.

## 5 DISCUSSION AND FUTURE DIRECTIONS

Our participants' reactions illustrate a diversity of attitudes toward the roles that AI services could play in their introspective practice. It was clear that currently available digital products fall short of supporting their needs. Our findings show that proposing concepts that exhibit potentially controversial design qualities can be productive in getting a better grasp on the design space. Through investigating concepts that appeared contentious, we were able to develop a sensibility for understanding where people's boundaries exist, what design qualities 'cross' them, and how people have different

viewpoints on social acceptability. Next, we present a summary of our key findings and then propose opportunities they suggest for future HCI research and design practice.

## 5.1 Summary of findings

*Co-Creation with new introspective Resources* (Theme 1) revealed that participants valued having more agency over what and how data is collected and that these acts were often perceived as opportunities to prompt introspection. Participants appreciated the ongoing dialogue about their data and valued the qualities of IAI services that invited them to be creative with those accounts. They also appreciated the IAI services that made their data malleable and tangible through both direct and implicit forms of interaction. Finally, participants valued tweaking, exaggerating, and re-contextualizing their digital profiles with the help of AI tools, which could lead to co-creative cycles of manifesting, discovering, and contemplating new aspects of their identity.

*New perspectives and confrontations of self* (Theme 2) located a nuanced tension concerning how an IAI presents information to the user. For participants, it was highly desired for an IAI to reveal patterns and introduce new perspectives on one's life; however, it was perceived as unacceptable to be told what these patterns might definitively mean. Participants wanted to have the opportunity to be more involved in an IAI's decision-making process and wanted options for 'opting in' and 'tuning' the degree to which an IAI service leveraged machine analysis for introspection. Surprisingly, participants also largely appreciated when the IAI exhibited bold and cheeky attitudes. These qualities were perceived as productive in pushing them out of their comfort zones, diverting from the status quo affirmational aims of most commercial applications. Yet, the more seasoned introspective practitioners expressed concern for less experienced users who could fall prey to technologically determined "self-fulfilling prophecies," marking areas that must be handled carefully.

*Desire for presence, scalability and new rituals* (Theme 3) showed that some participants envisioned longer-term human-technology relations with the IAI services where the contemplative rewards could deepen over time. They expressed enthusiasm for tools that could help them to explore temporal interrelations among their past-, present- and future selves. They also gravitated towards IAI services that were embodied in physical devices for their intentional, solitary focus on specific methods. Here, persistent, self-contained, safeguarded experiences were envisioned to nurture new and novel introspective rituals—they were often described as desired alternatives to smartphone apps that dominate the current design space. Lastly, participants expressed desires to socially share resources generated by IAI services as well as participate more directly in the development of IAI models that extend to broader communities of practice – both of which highlight key needs not well supported by currently available AI applications and services.

## 5.2 Extending co-creation as an IAI-enabled practice

Our findings revealed a desire for co-creative approaches that ranged from *ephemeral* to *labor intensive*, from *open-ended* to *guided* forms of interaction. Participants reacted highly positively

to services that created new resources which augmented existing introspective practices, such as expressive journaling, vision boarding, dream journaling, or mapping personal biases. Our study hints at various desires and needs if future introspective services were to include such strategies of co-creating resources with and through AI tools.

One strategy of co-creation with an IAI model focuses on *explicitly* prompting active data collection from end users. Participants perceived value in such 'hands-on' forms of interaction with the IAI to generate new resources, which they speculated could support anticipation and validate the labor people put into collecting data. That this type of interaction design could enable them to play a direct role in training their IAI model, even if it was uncertain how precisely these changes would be manifested, was nonetheless viewed positively. Participants also viewed more subtle acts of correcting and training their IAI model as another more *implicit* form of co-creative interaction with an IAI application. These interactions were perceived to empower users to augment their digital profiles to better understand how it reflected their self—even if this included forgotten or overlooked aspects of their behavior—or how the IAI model viewed them—which could mean radically augmenting different parameters to explore potentially exaggerated depictions of their self-image.

Taken together, our findings indicate explicit and implicit forms of interaction could offer distinct, yet potentially complementary, strategies to co-creatively engage with AI models and arrive at introspective resources. This initial validation is encouraging and suggests opportunities for future work in HCI.

Clearly, more research is needed to better understand how to practically design (i) AI techniques through which prompts are generated for individual users to collect new data for co-creation and (ii) the processes through which new generative introspective resources could be created that bring value to end users. As a pragmatic next step, future research could focus on developing, training, and trialing explicit Introspective AI prompt generators. As user would collect novel introspective data based on specific prompts, they could connect such accounts to more open-ended generative AI applications (e.g., DALL-E outpainting [110] or Artbreeder [130]) as productive starting points. Equally, there is a need to explore how co-creative implicit interaction can be mobilized through IAI applications. A practical next step could involve combining readily available forms of user data with AI sentiment analysis and subsequently using the outcome of the data and AI-inferred emotional valence to generate an obscured assessment, inspiring users to be contemplative while initially avoiding explicitly stating how the resource was created. For example, conducting a sentiment analysis on a one-week log of a user's text messages and using an OpenAI API (e.g., GPT-3 [30]) to generate a poetic statement on the inferred emotional climate of their last week. The user could then assess how 'correct' the statement feels and be revealed key attributes of the analysis that shaped the AI in generating it. In both cases, these practical next steps are intended not to be conclusive but rather generative of new research inquiries – such as the iterative development of technology probes [55] that could catalyze future research beyond speculative envisioning and into design practice-led innovation and empirical field studies of people's lived-with experiences with IAI applications.

### 5.3 Revealing patterns and new perspectives while balancing confrontation

Our findings reveal that there is a nuanced tension concerning how an IAI presents information to the user. For participants, it was highly desirable for an IAI to reveal patterns and introduce new perspectives on one's life; however, it was largely perceived as unacceptable to be told what these patterns definitively mean. This finding resonates with recent HCI research that has emphasized the value of AI uncertainty and unpredictability as design resources (e.g., [10, 54]) and foregrounds the importance of maintaining interpretable perspectives and situated sense-making [11]. In extending this work, our research suggests the need to explore and design IAI services that adopt explanation strategies that make space for users to interpret, analyze, and conclude what is laid out before them on their own terms. Design initiatives leveraging machine analysis for introspection should enable the user to opt into whether or not they want it. One compromise could be including a second interaction layer or an added lens that distinctly visualizes the inferences of data the machine draws upon, thus offering more insight into the inner workings of the IAI service and how patterns or new perspectives surfaced.

Many of the more advanced introspective practitioners identified that they would approach the inferences that an IAI might make as food for thought but not a representation of reality—resources in a similar light to horoscopes or fortune cookies. Yet, they were cautionary of the real possible dangers for less experienced practitioners who could become disillusioned with false images of themselves. This implication makes clear that IAI services will need to be able to grow and scale over time depending on the competencies of each user. At a minimum, this form of service should be tailored to an advanced group of practitioners, as well as designed in interactive overlays to remind less experienced users of their personal authority over these tools, which could fade away over time as their expertise grows.

### 5.4 Balancing curatorial control and autonomy

In our analysis, we were particularly fascinated by the responses to the role of IAI as an active curatorial agent. We discovered an interesting tension where balance needed to be found between the IAI holding users to account, yet without impeding their ability to have agency over their introspective resources. In the context of *Vision Shrine*, this meant enabling the user to be in control without losing the key functionalities of the device: playfully confronting people with their actions — or self-discrepancies — and holding them accountable in a way that changed over time.

These findings present several opportunities for future research. The implicit ludic design qualities expressed by *Vision Shrine* were well received by participants as potentially helping to mitigate the tension mentioned above—especially when combined, they were seen to evoke a light-hearted emotional valence that avoided being paternal or displeasing. Caraban et al. [19] describe “friction nudges” as interventions that disguise their intrusiveness yet uphold their potential to change people's behavior. Our findings suggest a promising opportunity for future research to test if the balance between AI-enacted intrusion and ludic interaction design can create more effective friction nudges or “things with attitude” [63].

One could imagine that people could be able to modulate an application like *Vision Shrines*' level of intrusiveness, relational to different levels of AI involvement. Participants also generated other design alternatives that often indicated desires for even more control. This included ‘pinning’ envisioned goals that they did not want to be augmented while enabling other unanchored elements to continue evolving through the IAI's ongoing actions. Or modulating the overall pacing of the IAI to slow down its actions and make space for reflecting on the contents of one's *Vision Board* from different points in time in the past. These insights illustrate different tactics to counterbalance the IAI, suggesting an opportunity for future research to leverage a combination of *ludic* [39], *careful* [50], and *slow* [47] design qualities to explore how interaction techniques can be effective at enacting implicit IAI curatorial actions in ways that scale over time. Future research in this space will provide new insights into the temporal dimensions of implicit curatorial control, helping better understand the potentialities for this strategy to scale over time.

There is also an opportunity to explore how users can be extended more control over the degree to which an IAI may intervene by, for example, integrating an adjustable setting that lets a user tune to their personal preference. Indeed, such added control could be scaffolded by the user as they become more adept at practicing introspection with their IAI model. Perhaps at first leveraging a high degree of control, but eventually ceding more curatorial autonomy if they developed a ‘take-it or leave-it’ sensibility. Yet, these processes of modulating curatorial control among users and their IAI will likely be complex. Balancing these two concerns harmoniously can be difficult and unpredictable, and more design exemplars are needed to illustrate effective interaction trajectories. Future work in this area can extend growing HCI research that investigates factors shaping how and when users are willing to cede agency to autonomous systems in light of the value that can emerge if control is ceded in a meaningful way [33, 39, 49, 76, 82].

### 5.5 Preserving and exploring temporal connections among a past, present, and future self

Many participants saw value in the co-created IAI archives of data and resources that they would develop over time and expressed their enthusiasm for tools that helped them to explore interconnections among their past-, present- and future selves. Such temporally rich interactions were most discussed around *Mind Probes*, where a user would record their current understanding of an emotion and, over time, be able to explore, compare and reflect on these representations. While the concept of resurfacing memories is present on consumer devices through machine vision analysis of photo archives (e.g., *Memories* on iOS) and has been explored in HCI through research on *slow technology* [47, 76] and *reminiscence* [22, 94], it is not clear how this may be leveraged in the context of highly personal affective data.

This poses new questions for the design of Introspective AI systems that aim to promote and facilitate temporal interactions: How can connections among a past, present, and future self be thoughtfully integrated into the core interaction design of Introspective AI services? If users gradually enrich their archives with temporally relevant data, how would those different stages of self be brought

into context? How might prompts be designed to bridge different time dimensions into specific themes relevant to one's life stage and life history?

While our study takes a first step in this direction, these findings and provocations present a clear opportunity for future research. Moving forwards, it is important to develop an understanding of how an IAI might detect when it is an appropriate time to surface an introspective historical resource (such as a journal entry), and through what metrics it might use to determine which entry is relevant, significant, and appropriate to the lived present.

## 5.6 Exploring IAI as a catalyst for human-to-human connection

We were surprised by participants' desires for *social* or *communal* dimensions to be extended to the IAI services, which were motivated by different and distinct needs. At the forefront, there was a desire to share and receive the generated output of the IAI services as a way of reinforcing accountability and learning about one's self and loved ones. There is a need for future research to explore how such highly personalized resources could be made intelligible to a broader social audience, such as friends and family, or potentially disclosed to trained professionals. Yet, this process will need to be handled carefully, and the user ought to be extended a degree of editorial control over what elements of their IAI are viewable and accessible to others. Research has shown how easy it is to overwrite personal beliefs by comparing one's self to others (i.e., *herd instinct bias* [93]); thus, selecting suitable ways of connecting and nudging people is key in building safeguarded experiences [19]. In stark contrast to other quantified self-applications, Gouveia et al. [44] proposed a strategy of comparison between users that are sharing similar goals (e.g., seeing someone else's progress in a running app), thus shrinking the margins that people could compare themselves to unrealistic standards that would pressure them. Future research is needed to explore the degree to which similar strategies would be viable if the IAI design space extended toward users sharing and comparing their introspective profiles. Clearly, even if well intended, there are dangers that this approach could lead to users tweaking their profiles to trend toward being more homogenous or have starker contrasts in comparison to others.

Prior research has shown that personal data that is valued for its reflective potential on an individual level may find a contentious place when integrated into a social archive and preserved over time [37, 45, 46, 75]. Participants also pushed back against the idea of IAI services emerging in the current e-health context. They desired to have a more active and agentic role within a community of stakeholders in the development of the IAI design space and the models shaping it. While likely a longer-term consideration, these findings indicate there is potential for emergent technologies that securely and collaboratively distribute power among various stakeholders, such as decentralized autonomous organizations [66] or data intermediary organizations [6], which could be viable alternatives for producing human-centric IAI infrastructure.

## 6 CONCLUSION

Extending nascent design research, we created four video-based design proposals to explore the emerging Introspective AI design

space. By probing on different, at times contentious, design qualities, we elicited complementary and contrasting attitudes toward how AI might offer new resources for supporting people's introspective practices. Our work opens possibilities for Introspective AI applications that, with the appropriate precautions, have the potential to diversify and expand this emerging design space based on new insights into what people desire and where boundaries of social acceptability exist. Our findings do raise several issues about the social appropriateness of new Introspective AI technology. We are currently developing technology probes based on the proposed opportunity areas to critically explore such issues. While our sample of participants aimed to be generative and varied in terms of age, occupation, and experience with introspection, they had a predominantly Western background. This limitation suggests an opportunity for future work to focus on participants from different cultures to see if and how attitudes may change. Ultimately, we hope this study inspires future research into how technologies could be designed to support and extend people's introspective practices in more valuable and value oriented ways.

## ACKNOWLEDGMENTS

This research took place in the Greater Vancouver area in Canada on the unceded traditional territories of the Coast Salish peoples of the Katzie, Kwantlen, Kwikwetlem (kwikwəłəm), Qayqayt, Musqueam (xwməθkwəyəm), and numerous Stó:lō Nations. This research is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Social Sciences and Humanities Research Council of Canada (SSHRC), and the Canada Foundation for Innovation (CFI). We thank our participants for generously sharing their experiences with us and Aamir Ali, Chiara Schmitt, Chiara Ferrari, Julian Goto, and Vanessa Montoya for their assistance on this project. We also thank the anonymous reviewers for their highly constructive feedback, which helped improve the quality of this paper.

## REFERENCES

- [1] 1-Life Inc. 2020. MertiLife: A vehicle for lifestyle change. Retrieved from <https://apps.apple.com/ca/app/metrilife/id1462361987>
- [2] Saleema Amershi, Maya Cakmak, William Bradley Knox, and Todd Kulesza. 2014. Power to the People: The Role of Humans in Interactive Machine Learning. *AI Magazine* 35, 4: 105–120. <https://doi.org/10.1609/aimag.v35i4.2513>
- [3] Saleema Amershi, James Fogarty, and Daniel Weld. 2012. Regroup: interactive machine learning for on-demand group creation in social networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 21–30. Retrieved August 18, 2021 from <http://doi.org/10.1145/2207676.2207680>
- [4] James Auger. 2013. Speculative design: crafting the speculation. *Digital Creativity* 24, 1: 11–35. <https://doi.org/10.1080/14626268.2013.767276>
- [5] Ciano Aydin. 2021. *Extimate Technology: Self-Formation in a Technological World*. Taylor & Francis. <https://doi.org/10.4324/9781003139409>
- [6] Samuel Barnett, Nico Brand, William Odom, and Katlyn Andres. Exploring Data Intermediaries as Infrastructure for a Human-Centric Data Economy: Speculations & Critical Reflections. In *Proceedings of the Nordic conference on human-computer interaction: participative computing for sustainable futures*, 1–20.
- [7] Russell W. Belk. 1988. Possessions and the extended self. *Journal of consumer research* 15, 2: 139–168.
- [8] Russell W. Belk. 2013. Extended self in a digital world. *Journal of consumer research* 40, 3: 477–500.
- [9] Jesse Josua Benjamin, Arne Berger, Nick Merrill, and James Pierce. 2021. Machine Learning Uncertainty as a Design Material: A Post-Phenomenological Inquiry. *arXiv preprint arXiv:2101.04035*.
- [10] 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*. Association for Computing Machinery, New York, NY, USA, 1–14. Retrieved August 20, 2021 from <https://doi.org/10.1145/3411764.3445481>
- [11] Jesse Josua Benjamin, Christoph Kinkeldey, Claudia Müller-Birn, Tim Korjakow, and Eva-Maria Herbst. 2022. Explanation Strategies as an Empirical-Analytical Lens for Socio-Technical Contextualization of Machine Learning Interpretability. *Proceedings of the ACM on Human-Computer Interaction* 6, GROU: 39:1-39:25. <https://doi.org/10.1145/3492858>
  - [12] Julian Bleeker. 2009. *Design Fiction: A Short Essay on Design, Science, Fact and Fiction*. Retrieved April 5, 2017 from [http://drbfw5wfwjlxon.cloudfront.net/writing/DesignFiction\\_WebEdition.pdf](http://drbfw5wfwjlxon.cloudfront.net/writing/DesignFiction_WebEdition.pdf)
  - [13] Mark Blythe. 2014. Research Through Design Fiction: Narrative in Real and Imaginary Abstracts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*, 703–712. <https://doi.org/10.1145/2556288.2557098>
  - [14] Mark Blythe. 2017. Research Fiction: Storytelling, Plot and Design. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*, 5400–5411. <https://doi.org/10.1145/3025453.3026023>
  - [15] Renaud Bougueng Tcheneube, Jeffrey John Ens, and Philippe Pasquier. 2022. Calliope: A Co-creative Interface for Multi-Track Music Generation. In *Creativity and Cognition*, 608–611. <https://doi.org/10.1145/3527927.3535200>
  - [16] 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 (DIS '21)*, 1603–1618. <https://doi.org/10.1145/3461778.3462000>
  - [17] Lisa Burton and Jonathan Lent. 2016. The Use of Vision Boards as a Therapeutic Intervention. *Journal of Creativity in Mental Health* 11, 1: 52–65. <https://doi.org/10.1080/15401383.2015.1092901>
  - [18] Alex Byrne. 2005. Introspection. *Philosophical Topics* 33, 1: 79–104.
  - [19] Ana Caraban, Evangelos Karapanos, Daniel Gonçalves, and Pedro Campos. 2019. 23 Ways to Nudge: A Review of Technology-Mediated Nudging in Human-Computer Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15. <https://doi.org/10.1145/3290605.3300733>
  - [20] Baptiste Caramiaux and Sarah Fdili Alaoui. 2022. “Explorers of Unknown Planets”: Practices and Politics of Artificial Intelligence in Visual Arts. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2: 477:1-477:24. <https://doi.org/10.1145/3555578>
  - [21] Michelle Carney, Barron Webster, Irene Alvarado, Kyle Phillips, Noura Howell, Jordan Griffith, Jonas Jongejan, Amit Pitaru, and Alexander Chen. 2020. Teachable Machine: Approachable Web-Based Tool for Exploring Machine Learning Classification. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*, 1–8. <https://doi.org/10.1145/3334480.3382839>
  - [22] Amy Yo Sue Chen, William Odom, Ce Zhong, Henry Lin, and Tal Amram. 2019. Chronoscope: Designing Temporally Diverse Interactions with Personal Digital Photo Collections. In *Proceedings of the 2019 on Designing Interactive Systems Conference*, 799–812.
  - [23] Janet X. Chen, Francesco Vitale, and Joanna McGrenere. 2021. What Happens After Death? Using a Design Workbook to Understand User Expectations for Preparing their Data. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. Retrieved August 23, 2021 from <https://doi.org/10.1145/3411764.3445359>
  - [24] Victoria Clarke and Virginia Braun. 2014. Thematic Analysis. In *Encyclopedia of Critical Psychology*, Thomas Teo (ed.). Springer, New York, NY, 1947–1952. [https://doi.org/10.1007/978-1-4614-5583-7\\_311](https://doi.org/10.1007/978-1-4614-5583-7_311)
  - [25] Jemine Coarentin. 2019. Real-Time-Voice-Cloning. Retrieved from <https://github.com/Coarentin/Real-Time-Voice-Cloning>
  - [26] Paul Coulton, Joseph Galen Lindley, Miriam Sturdee, and Michael Stead. 2017. Design fiction as world building.
  - [27] Kate Crawford. 2021. *The Atlas of AI*. Yale University Press.
  - [28] Kate Crawford and Vladan Joler. 2018. Anatomy of an AI System. *AI NOW Institute*: 2018.
  - [29] Amber L. Cushing. 2013. “It’s stuff that speaks to me”: Exploring the characteristics of digital possessions. *Journal of the American Society for Information Science and Technology* 64, 8: 1723–1734.
  - [30] Robert Dale. 2021. GPT-3: What’s it good for? *Natural Language Engineering* 27, 1: 113–118.
  - [31] Daniel C. Dennett. 1976. Are dreams experiences? *The Philosophical Review* 85, 2: 151–171.
  - [32] Audrey Desjardins, Jeremy E. Viny, Cayla Key, and Nouela Johnston. 2019. Alternative Avenues for IoT: Designing with Non-Stereotypical Homes. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*, 1–13. <https://doi.org/10.1145/3290605.3300581>
  - [33] Laura Devendorf and Kimiko Ryokai. 2015. Being the Machine: Reconfiguring Agency and Control in Hybrid Fabrication. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*, 2477–2486. <https://doi.org/10.1145/2702123.2702547>
  - [34] 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. Retrieved from [http://dl.acm.org/citation.cfm?id\\$=33025739](http://dl.acm.org/citation.cfm?id$=33025739)
  - [35] Carolyn Ellis. 1991. Sociological introspection and emotional experience. *Symbolic interaction* 14, 1: 23–50.
  - [36] Chris Elsdén, David Chatting, Abigail C. Durrant, Andrew Garbett, Bettina Nissen, John Vines, and David S. Kirk. 2017. On Speculative Enactments. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*, 5386–5399. <https://doi.org/10.1145/3025453.3025503>
  - [37] Chris Elsdén, David S. Kirk, and Abigail C. Durrant. 2016. A Quantified Past: Toward Design for Remembering With Personal Informatics. *Human-Computer Interaction* 31, 6: 518–557. <https://doi.org/10.1080/07370024.2015.1093422>
  - [38] Shi Feng and Jordan Boyd-Graber. 2019. What can AI do for me? evaluating machine learning interpretations in cooperative play. In *Proceedings of the 24th International Conference on Intelligent User Interfaces (IUI '19)*, 229–239. <https://doi.org/10.1145/3301275.3302265>
  - [39] William W. Gaver, John Bowers, Andrew Boucher, Hans Gellerson, Sarah Pennington, Albrecht Schmidt, Anthony Steed, Nicholas Villars, and Brendan Walker. 2004. The Drift Table: Designing for Ludic Engagement. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems (CHI EA '04)*, 885–900. <https://doi.org/10.1145/985921.985947>
  - [40] Maliheh Ghajargar and Jaffrey Bardzell. 2022. Making AI Understandable by Making it Tangible: Exploring the Design Space with Ten Concept Cards. In *The 34th Australian Conference on Human-Computer Interaction (OzCHI'22) Canberra, Australia 29 November–2 December. 2022*.
  - [41] Elisa Giacardi and Johan Redström. 2020. Technology and More-Than-Human Design. *Design Issues* 36, 4: 33–44. [https://doi.org/10.1162/desi\\_a\\_00612](https://doi.org/10.1162/desi_a_00612)
  - [42] Dorota Glowacka, Tuukka Ruotsalo, Ksenia Konuyshkova, kumaripaba Athukorala, Samuel Kaski, and Giulio Jacucci. 2013. Directing exploratory search: reinforcement learning from user interactions with keywords. In *Proceedings of the 2013 international conference on Intelligent user interfaces (IUI '13)*, 117–128. <https://doi.org/10.1145/2449396.2449413>
  - [43] Stephen J. Gould. 1995. Researcher introspection as a method in consumer research: Applications, issues, and implications. *Journal of consumer research* 21, 4: 719–722.
  - [44] Rúben Gouveia, Fábio Pereira, Evangelos Karapanos, Sean A. Munson, and Marc Hassenzahl. 2016. Exploring the design space of glanceable feedback for physical activity trackers. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 144–155. <https://doi.org/10.1145/2971648.2971754>
  - [45] Rebecca Gulotta, William Odom, Haakon Faste, and Jodi Forlizzi. 2014. Legacy in the Age of the Internet: Reflections on How Interactive Systems Shape How We Are Remembered. In *Proceedings of the 2014 Conference on Designing Interactive Systems (DIS '14)*, 975–984. <https://doi.org/10.1145/2598510.2598579>
  - [46] Rebecca Gulotta, Alex Sciuto, Aisling Kelliher, and Jodi Forlizzi. 2015. Curatorial agents: How systems shape our understanding of personal and familial digital information. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3453–3462. Retrieved from [http://dl.acm.org/citation.cfm?id\\$=2702297](http://dl.acm.org/citation.cfm?id$=2702297)
  - [47] Lars Hallnäs and Johan Redström. 2001. Slow technology—designing for reflection. *Personal and ubiquitous computing* 5, 3: 201–212.
  - [48] Sabrina Hauser, Johan Redström, and Heather Wiltse. 2021. The widening rift between aesthetics and ethics in the design of computational things. *AI & SOCIETY*. <https://doi.org/10.1007/s00146-021-01279-w>
  - [49] John Helmes, Alex S. Taylor, Xiang Cao, Kristina Höök, Peter Schmitt, and Nicolas Villar. 2011. Rudiments 1, 2 & 3: design speculations on autonomy. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, 145–152. Retrieved from [http://dl.acm.org/citation.cfm?id\\$=1935730](http://dl.acm.org/citation.cfm?id$=1935730)
  - [50] Karey Helms. 2020. Careful Design: Implicit Interactions with Care, Taboo, and Humor. In *Companion Publication of the 2020 ACM Designing Interactive Systems Conference (DIS' 20 Companion)*, 515–519. <https://doi.org/10.1145/3393914.3395827>
  - [51] E. Tory Higgins. 1989. Self-Discrepancy Theory: What Patterns of Self-Beliefs Cause People to Suffer? In *Advances in Experimental Social Psychology*. Elsevier, 93–136. [https://doi.org/10.1016/S0065-2601\(08\)60306-8](https://doi.org/10.1016/S0065-2601(08)60306-8)
  - [52] Tad Hirsch, Kritzia Merced, Shrikanth Narayanan, Zac E. Imel, and David C. Atkins. 2017. Designing Contestability: Interaction Design, Machine Learning, and Mental Health. In *Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17)*, 95–99. <https://doi.org/10.1145/3064663.3064703>
  - [53] Tad Hirsch, Christina Soma, Kritzia Merced, Patty Kuo, Aaron Dembe, Derek D. Caperton, David C. Atkins, and Zac E. Imel. 2018. “It’s hard to argue with a computer”: Investigating Psychotherapists’ Attitudes towards Automated Evaluation. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*, 559–571. <https://doi.org/10.1145/3196709.3196776>
  - [54] Stacy Hsueh, Sarah Fdili Alaoui, and Wendy E. Mackay. 2019. Understanding Kinesthetic Creativity in Dance. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*, 1–12. <https://doi.org/10.1145/>



- 3290605.3300741
- [55] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B. Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, and others. 2003. Technology probes: inspiring design for and with families. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 17–24. Retrieved September 19, 2015 from [http://dl.acm.org/citation.cfm?id=\\$642616](http://dl.acm.org/citation.cfm?id=$642616)
- [56] Don Ihde and Lambros Malafouris. 2019. Homo faber Revisited: Postphenomenology and Material Engagement Theory. *Philosophy & Technology* 32, 2: 195–214. <https://doi.org/10.1007/s13347-018-0321-7>
- [57] Ye Jia, Yu Zhang, Ron J. Weiss, Quan Wang, Jonathan Shen, Fei Ren, Zhifeng Chen, Patrick Nguyen, Ruoming Pang, Ignacio Lopez Moreno, and Yonghui Wu. 2018. Transfer learning from speaker verification to multispeaker text-to-speech synthesis. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18)*, 4485–4495.
- [58] Vera Khovanskaya, Eric P.S. Baumer, Dan Cosley, Stephen Volda, and Geri Gay. 2013. "Everybody knows what you're doing": a critical design approach to personal informatics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*, 3403–3412. <https://doi.org/10.1145/2470654.2466467>
- [59] Asle H. Kiran and Peter-Paul Verbeek. 2010. Trusting Our Selves to Technology. *Knowledge, Technology & Policy* 23, 3: 409–427. <https://doi.org/10.1007/s12130-010-9123-7>
- [60] Alexandra Kitson, Thecla Schiphorst, and Bernhard E. Riecke. 2018. Are You Dreaming? A Phenomenological Study on Understanding Lucid Dreams as a Tool for Introspection in Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–12. Retrieved February 7, 2021 from <https://doi.org/10.1145/3173574.3173917>
- [61] Mike Kuniavsky, Elizabeth Churchill, and Molly Wright Steenson. 2017. *The 2017 aaai spring symposium series technical reports: Designing the user experience of machine learning systems*. Technical Report SS-17-04. Palo Alto, California.
- [62] Himabindu Lakkaraju, Ece Kamar, Rich Caruana, and Jure Leskovec. 2019. Faithful and Customizable Explanations of Black Box Models. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (AIES '19)*, 131–138. <https://doi.org/10.1145/3306618.3314229>
- [63] Matthias Laschke, Marc Hassenzahl, and Sarah Diefenbach. Things with attitude: Transformational Products.
- [64] 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*. Association for Computing Machinery, New York, NY, USA, 1–13. Retrieved August 18, 2021 from <http://doi.org/10.1145/3313831.3376792>
- [65] Jonas Löwgren. 2004. Animated use sketches as design representations. *Interactions* 11, 6: 22–27. <https://doi.org/10.1145/1029036.1029048>
- [66] Caitlin Lustig. 2019. Intersecting Imaginaries: Visions of Decentralized Autonomous Systems. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW: 210:1–210:27. <https://doi.org/10.1145/3359312>
- [67] Trieuvy Luu, Martijn van den Broeck, and Marie Louise Juul Søndergaard. 2018. Data economy: interweaving storytelling and world building in design fiction. In *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*, 771–786. <https://doi.org/10.1145/3240167.3240270>
- [68] Wendy E. Mackay, Anne V. Ratzer, and Paul Janecsek. 2000. Video artifacts for design: bridging the Gap between abstraction and detail. In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques (DIS '00)*, 72–82. <https://doi.org/10.1145/347642.347666>
- [69] 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 (DIS '19)*, 1083–1088. <https://doi.org/10.1145/3322276.3322351>
- [70] Clara Mancini, Yvonne Rogers, Arosha K. Bandara, Tony Coe, Lukasz Jedrzejczyk, Adam N. Joinson, Blaine A. Price, Keerthi Thomas, and Bashar Nuseibeh. 2010. Contravision: exploring users' reactions to futuristic technology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 153–162. Retrieved August 16, 2021 from <http://doi.org/10.1145/1753326.1753350>
- [71] Matthew B. Miles and A. Michael Huberman. 1985. *Qualitative data analysis*. Sage Newbury Park, CA. Retrieved January 11, 2017 from <http://researchtalk.com/wp-content/uploads/2014/01/Miles-Huberman-Saldana-Drawing-and-Verifying-Conclusions.pdf>
- [72] Dave Murray-Rust, Johanna Nicenboim, and Dan Lockton. 2022. Metaphors for designers working with AI. *DRS Biennial Conference Series*. Retrieved from <https://dl.designresearchsociety.org/drs-conference-papers/drs2022/researchpapers/237>
- [73] Johanna 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 (DIS' 20 Companion)*, 397–400. <https://doi.org/10.1145/3393914.3395912>
- [74] Warren T. Norman. 1963. Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The Journal of Abnormal and Social Psychology* 66, 6: 574–583. <https://doi.org/10.1037/h0040291>
- [75] William Odom, Richard Banks, David Kirk, Richard Harper, Siân Lindley, and Abigail Sellen. 2012. Technology Heirlooms?: Considerations for Passing Down and Inheriting Digital Materials. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*, 337–346. <https://doi.org/10.1145/2207676.2207723>
- [76] William T. Odom, Abigail J. Sellen, Richard Banks, David S. Kirk, Tim Regan, Mark Selby, Jodi L. Forlizzi, and John Zimmerman. 2014. Designing for Slowness, Anticipation and Re-visitation: A Long Term Field Study of the Photobox. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*, 1961–1970. <https://doi.org/10.1145/2556288.2557178>
- [77] William Odom, John Zimmerman, Scott Davidoff, Jodi Forlizzi, Anind K. Dey, and Min Kyung Lee. 2012. A Fieldwork of the Future with User Enactments. In *Proceedings of the Designing Interactive Systems Conference (DIS '12)*, 338–347. <https://doi.org/10.1145/2317956.2318008>
- [78] William Odom, John Zimmerman, and Jodi Forlizzi. 2014. Placelessness, Spacelessness, and Formlessness: Experiential Qualities of Virtual Possessions. In *Proceedings of the 2014 Conference on Designing Interactive Systems (DIS '14)*, 985–994. <https://doi.org/10.1145/2598510.2598577>
- [79] Josef Perner, Daniela Kloof, and Elisabeth Stöttinger. 2007. Introspection & remembering. *Synthese* 159, 2: 253–270.
- [80] Alicia Nortje Ph.D. 2020. What Is Cognitive Bias? 7 Examples & Resources (Incl. Codex). *PositivePsychology.com*. Retrieved December 7, 2022 from <https://positivepsychology.com/cognitive-biases/>
- [81] Franziska Louise Pilling and Paul Coulton. 2020. What's it like to be Alexa? An exploration of Artificial Intelligence as a Material for Design. *DRS Biennial Conference Series*. Retrieved from <https://dl.designresearchsociety.org/drs-conference-papers/drs2020/researchpapers/53>
- [82] Zachary Pousman, Mario Romero, Adam Smith, and Michael Mateas. 2008. Living with tableau machine: a longitudinal investigation of a curious domestic intelligence. In *Proceedings of the 10th international conference on Ubiquitous computing*, 370–379. Retrieved from [http://dl.acm.org/citation.cfm?id=\\$1409685](http://dl.acm.org/citation.cfm?id=$1409685)
- [83] S. R. F. Price. 1986. The Future of Dreams: From Freud to Artemidorus. *Past & Present*, 113: 3–37.
- [84] Larissa Pschetz, Krakaek Pothong, and Chris Speed. 2019. Autonomous Distributed Energy Systems: Problematising the Invisible through Design, Drama and Deliberation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. Retrieved August 16, 2021 from <http://doi.org/10.1145/3290605.3300617>
- [85] Amon Rapp and Maurizio Tirassa. 2017. Know Thyself: A Theory of the Self for Personal Informatics. *Human-Computer Interaction* 32, 5–6: 335–380. <https://doi.org/10.1080/07370024.2017.1285704>
- [86] John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers. 2014. Personal Tracking As Lived Informatics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*, 1163–1172. <https://doi.org/10.1145/2556288.2557039>
- [87] Karin Ryding. 2020. The Silent Conversation: Designing for Introspection and Social Play in Art Museums. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–10. Retrieved August 16, 2021 from <http://doi.org/10.1145/3313831.3376357>
- [88] Eric Schwitzgebel. 2012. Introspection, what? *Introspection and consciousness*: 29–48.
- [89] Sydney Shoemaker. 1986. Introspection and the Self. *Midwest Studies in Philosophy* 10: 101–120.
- [90] Paul J. Silvia and Guido H. E. Gendolla. 2001. On Introspection and Self-Perception: Does Self-Focused Attention Enable Accurate Self-Knowledge? *Review of General Psychology* 5, 3: 241–269. <https://doi.org/10.1037/1089-2680.5.3.241>
- [91] Alison Smith-Renner, Ron Fan, Melissa Birchfield, Tongshuang Wu, Jordan Boyd-Graber, Daniel S. Weld, and Leah Findlater. 2020. No Explainability without Accountability: An Empirical Study of Explanations and Feedback in Interactive ML. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. Retrieved August 16, 2021 from <http://doi.org/10.1145/3313831.3376624>
- [92] 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 (DIS '18)*, 869–880. <https://doi.org/10.1145/3196709.3196766>
- [93] Wilfred Trotter. 1908. Herd instinct and its bearing on the psychology of civilised man. *The Sociological Review* 1, 3: 227–248.
- [94] Wenn-Chieh Tsai, Po-Hao Wang, Hung-Chi Lee, Rung-Huei Liang, and Jane Hsu. 2014. The reflexive printer: toward making sense of perceived drawbacks in technology-mediated reminiscence. In *Proceedings of the 2014 conference*

- on *Designing interactive systems*, 995–1004. Retrieved from [http://dl.acm.org/citation.cfm?id=\\$2598589](http://dl.acm.org/citation.cfm?id=$2598589)
- [95] Robert Van Gulick. 2000. Inward and upward: reflection, introspection, and self-awareness. *Philosophical Topics* 28, 2: 275–305.
- [96] Peter-Paul Verbeek. 2010. *What things do: Philosophical reflections on technology, agency, and design*. Penn State Press. Retrieved January 20, 2017 from [https://books.google.ca/books?hl=\\$en&lr=\\$&id=\\$vURh8gy8nPAC&oi=\\$fnd&pg=\\$P11&dq=\\$What+things+do:+Philosophical+reflections+on+technology&ots=\\$cLJ\\_-QzIR4&sig=\\$04v4q9OlmELREhWRUAXo-c-W\\_gs](https://books.google.ca/books?hl=$en&lr=$&id=$vURh8gy8nPAC&oi=$fnd&pg=$P11&dq=$What+things+do:+Philosophical+reflections+on+technology&ots=$cLJ_-QzIR4&sig=$04v4q9OlmELREhWRUAXo-c-W_gs)
- [97] John Vines, Mark Blythe, Stephen Lindsay, Paul Dunphy, Andrew Monk, and Patrick Olivier. 2012. Questionable concepts: critique as resource for designing with eighty somethings. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1169–1178. Retrieved August 11, 2021 from <http://doi.org/10.1145/2207676.2208567>
- [98] Francesco Vitale, William Odom, and Joanna McGrenere. 2019. Keeping and Discarding Personal Data: Exploring a Design Space. In *Proceedings of the 2019 on Designing Interactive Systems Conference (DIS '19)*, 1463–1477. <https://doi.org/10.1145/3322276.3322300>
- [99] Matthew Walker. 2017. *Why we sleep: Unlocking the power of sleep and dreams*. Simon and Schuster.
- [100] Danding Wang, Qian Yang, Ashraf Abdul, and Brian Y. Lim. 2019. Designing Theory-Driven User-Centric Explainable AI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–15. Retrieved August 18, 2021 from <https://doi.org/10.1145/3290605.3300831>
- [101] Richmond Y. Wong and Deirdre K. Mulligan. 2016. When a Product Is Still Fictional: Anticipating and Speculating Futures through Concept Videos. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*, 121–133. <https://doi.org/10.1145/2901790.2901801>
- [102] Qian Yang. 2018. Machine learning as a UX design material: How can we imagine beyond automation, recommenders, and reminders? In *AAAI Spring Symposia*.
- [103] 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 (CHI '18)*, 1–11. <https://doi.org/10.1145/3173574.3173704>
- [104] Qian Yang, Jina Suh, Nan-Chen Chen, and Gonzalo Ramos. 2018. Grounding Interactive Machine Learning Tool Design in How Non-Experts Actually Build Models. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*, 573–584. <https://doi.org/10.1145/3196709.3196729>
- [105] J. Zimmerman. 2005. Video sketches: exploring pervasive computing interaction designs. *IEEE Pervasive Computing* 4, 4: 91–94. <https://doi.org/10.1109/MPRV.2005.91>
- [106] Shoshana Zuboff. 2015. Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology* 30, 1: 75–89. <https://doi.org/10.1057/jit.2015.5>
- [107] 2015. The Reason Vision Boards Work and How to Make One. *HuffPost*. Retrieved December 7, 2022 from [https://www.huffpost.com/entry/the-scientific-reason-why\\_b\\_6392274](https://www.huffpost.com/entry/the-scientific-reason-why_b_6392274)
- [108] 2021. What is Natural Language Processing? Retrieved December 7, 2022 from <https://www.ibm.com/cloud/learn/natural-language-processing>
- [109] 2022. Curatorial A(i)gents. *metaLAB (at) Harvard & FU Berlin*. Retrieved December 7, 2022 from <https://mlml.io/p/curatorial-aigents/>
- [110] 2022. DALL-E: Introducing Outpainting. *OpenAI*. Retrieved December 12, 2022 from <https://openai.com/blog/dall-e-introducing-outpainting/>
- [111] 2022. IBM Watson Speech to Text - Overview. Retrieved December 7, 2022 from <https://www.ibm.com/cloud/watson-speech-to-text>
- [112] Replika. *replika.ai*. Retrieved August 18, 2021 from <https://replika.ai>
- [113] OpenAI. *OpenAI*. Retrieved July 29, 2022 from <https://openai.com/>
- [114] Reflectly - A Journal for Happiness. Retrieved April 21, 2021 from <https://reflectly.app/>
- [115] The Superhuman App - Self Reliant And In Control - from Zenegant. *Zenegant*. Retrieved April 21, 2021 from <https://zenegant.io/zenegant-superhuman-strategic-thinking-self-reliant-control-decisions-emotions-interactions/>
- [116] Dreamapp. *Dreamapp*. Retrieved April 21, 2021 from <https://dreamapp.io>
- [117] The Artist's Way | Julia Cameron Live. Retrieved December 7, 2022 from <https://juliacameronlive.com/the-artists-way/>
- [118] What is art journaling. *Rachel Greig*. Retrieved December 7, 2022 from <https://rachelgreig.com/blog/what-is-art-journaling>
- [119] Capture - Your Dream Journal. *App Store*. Retrieved January 12, 2022 from <https://apps.apple.com/us/app/capture-your-dream-journal/id968737914>
- [120] RunwayML | Machine learning for creators. *RunwayML*. Retrieved from <https://runwayml.com>
- [121] Lobe | Machine Learning Made Easy. Retrieved December 29, 2021 from <https://www.lobe.ai/>
- [122] Generate Endless Musical Ideas with Live's MIDI Devices | Ableton. Retrieved December 7, 2022 from <https://www.ableton.com/en/blog/generate-endless-musical-ideas-lives-midi-devices/>
- [123] Endel - Personalized soundscapes to help you focus, relax, and sleep. Backed by neuroscience. *Endel - Personalized soundscapes to help you focus, relax, and sleep. Backed by neuroscience*. Retrieved December 7, 2022 from <https://endel.io/>
- [124] The Myers & Briggs Foundation - MBTI@Basics. Retrieved November 28, 2022 from [https://www.myersbriggs.org/my-mbti-personality-type/mbti-basics/home.htm?bhcp=\\$1](https://www.myersbriggs.org/my-mbti-personality-type/mbti-basics/home.htm?bhcp=$1)
- [125] IBM Watson | IBM. Retrieved from <https://www.ibm.com/watson>
- [126] The human insights operating system. Receptiviti. Retrieved January 6, 2022 from <https://www.receptiviti.com>
- [127] DALL-E 2. *OpenAI*. Retrieved December 7, 2022 from <https://openai.com/dall-e-2/>
- [128] Stable Diffusion Online. Retrieved December 7, 2022 from <https://stablediffusionweb.com/>
- [129] WOMBO Dream - AI Powered Artwork Tool. Retrieved December 30, 2021 from <https://www.wombo.art/>
- [130] Artbreeder. Retrieved December 12, 2022 from <https://www.artbreeder.com/>