Designing for Human-AI Interaction: Comparing Intermittent, Continuous, and Proactive Interactions for a Music Application

Anders Gammelgård-Larsen
anga@cs.aau.dk
Aalborg University
Aalborg, Denmark

Niels van Berkel
nielsvanberkel@cs.aau.dk
Aalborg University
Aalborg, Denmark

Mikael B. Skov
dubois@cs.aau.dk
Aalborg University
Aalborg, Denmark

Jesper Kjeldskov
jesper@cs.aau.dk
Aalborg University
Aalborg, Denmark

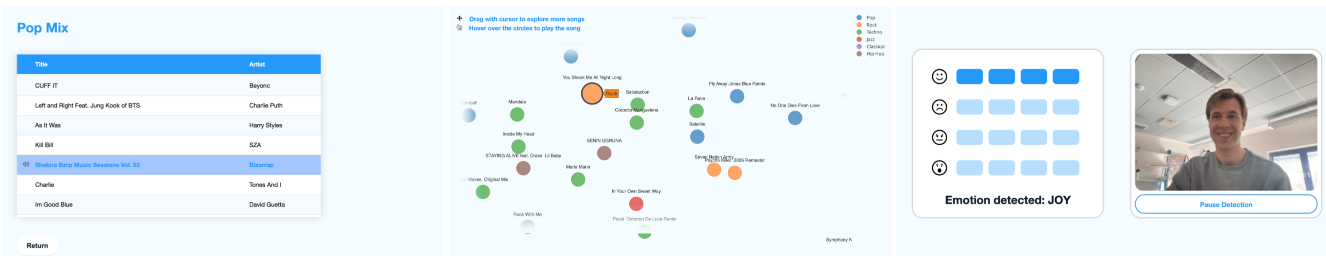

Figure 1: Three prototypes representing intermittent (left), continuous (centre), and proactive (right) human-AI interaction.

ABSTRACT

Designing effective and user-centred interactions between humans and AI systems poses fundamental challenges. The behaviour of AI systems is complex and uncertain, making it difficult to envision and craft optimal user experiences. Improved frameworks are needed to guide the design of human-AI interaction. In this paper, we develop and evaluate prototypes for a music application, representing three distinct paradigms of human-AI interaction: Intermittent, Continuous, and Proactive. Through qualitative user interviews with 12 participants, we compare the user experience across these prototypes, shedding light on potential challenges and opportunities for the paradigms represented. We found that the three prototypes exhibit distinct characteristics in terms of supported goals and user control. This case study contributes to a deeper understanding of the complexities involved in designing AI systems and offers insights for the development of more user-centred AI applications.

CCS CONCEPTS

• **Human-centered computing → Interaction paradigms**; *HCI theory, concepts and models*; *Empirical studies in HCI*.

ACM Reference Format:

Anders Gammelgård-Larsen, Niels van Berkel, Mikael B. Skov, and Jesper Kjeldskov. 2024. Designing for Human-AI Interaction: Comparing Intermittent, Continuous, and Proactive Interactions for a Music Application. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3613905.3650886

CHI EA '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0331-7/24/05
https://doi.org/10.1145/3613905.3650886

1 INTRODUCTION

Artificial Intelligence (AI) has made significant technological advancements, enabling a range of real-world applications across various domains. AI powers established product categories such as writing assistants, self-driving cars, and medical diagnosis systems, while also serving as a versatile design material for novel applications. However, along with the opportunities offered by AI, there are also many challenges for designers. Envisioning the to-be user experience and crafting the flow of interaction is problematic when the behaviour of most AI systems is complex and uncertain [25]. Therefore, improved frameworks for making sense of the different forms of human-AI interaction are required.

Multiple frameworks have been proposed to make sense of human-AI interactions (for example, [9, 10, 14]). While these frameworks provide a valuable way of conceptualising the interaction, few case studies have been conducted that directly compare different interaction paradigms. In this paper, we focus on a specific framework proposed by Van Berkel et al. [23]. This framework outlines three interaction paradigms that are distinctively different in terms of user control and initiative: *Intermittent*, *Continuous*, and *Proactive* interactions.

We present a case study that applies this framework to the design of a music application. The case study aims to investigate end-user implications and design considerations across prototypes representing the three different interaction paradigms. Our key findings show that the prototypes exhibit distinct characteristics in terms of supported goals and user control. The Intermittent prototype offers the most direct control and allows the user to effectively request specific songs. The Continuous prototype offers only limited control but supports the exploration of new music. With the Proactive prototype, users can listen to adaptive music, but it requires users to entrust the AI with full control.

The work's contributions extend beyond the specific context of music applications. Our research provides a deeper understanding of human-AI interaction paradigms, emphasising the importance of user control and initiative. Our findings offer insights and practical guidance for AI system designers.

2 RELATED WORK

2.1 Challenges for the design of AI systems

Designing responsible and ethical AI systems presents numerous challenges that necessitate careful consideration. These challenges include mitigating bias [12], protecting privacy [21], ensuring accountability [19], and integrating human empathy [18]. Design guidelines exist to ensure that these types of challenges are considered in the design of new AI products. However, AI is an ambiguous concept, referring to a wide range of techniques and applications (e.g., computer vision, natural language processing, big data analytics) [25, p.4]. Therefore, a designer might face challenges that are not considered by any guidelines, or a designer might find guidelines that do not apply to their specific challenges. Furthermore, intelligence as a technological material is often confused with intelligence as an experiential factor (i.e. a system that is developed to be 'intelligent' is not necessarily experienced as 'intelligent' by the end-user, and vice versa) [25, p.4]. Therefore, researchers should aim to understand the design challenges not only from a technical perspective but also in relation to how the AI is experienced.

According to Yang et al. many of the unique challenges in AI stem from varied degrees of *'capability uncertainty'* and *'output complexity'* [25]. An inherent property of evolving and adaptive AI systems is that designers cannot be certain how well the system will perform and which exact outputs it might generate during actual use. This results in fundamental challenges for designers to overcome when envisioning and prototyping human-AI interaction.

2.2 Guidelines for the design of AI systems

Prior works have proposed various guidelines that aim to support AI practitioners. In 1999, Horvitz proposed 12 principles for mixed-initiative user interfaces [10]. Generally, these early principles focused on how to effectively balance automated services and direct user control, with specific considerations such as the ideal timing of services and dealing with uncertainties. Horvitz's paper was highly influential for the emerging field of human-centred AI and inspired many further investigations into effective collaboration with intelligent agents.

More recently, in 2019, Amershi et al. synthesised over 150 recommendations from both academic and industry sources to produce a set of 18 generally applicable guidelines for designing human-AI interaction [1]. Many of these new guidelines touch upon the same ideas as discussed in Horvitz's paper from twenty years prior. However, with an additional focus on issues related to transparency, fairness, reliability and trust - reflecting temporary concerns in the use of AI technology. As evidence, major tech companies including Google [7], Microsoft [4] and IBM [3] have all documented a response to these issues within their official AI design principles.

However, Yang et al. argue that many of the guidelines for developing AI are applicable to any work within HCI and UX and that much of current research fails to distinguish between these and guidelines that address challenges unique to AI [25].

2.3 Making sense of Human-AI Interaction

Multiple frameworks have been proposed to classify the forms of interaction between people and intelligent systems. Methnani et al. discuss the varying levels of human control and summarise the different approaches as *Human-In-the-Loop*, *Human-On-the-Loop* and *Human-Out-of-the-loop*, where the human in the interaction with an AI system plays an integral role, supervisory role, or no role respectively [14]. Hinsen et al. identify at least nine characteristic dimensions (e.g., *freedom of action* and *reciprocal engagement*) and use these to distinguish five types of agents (e.g., *Guardian angel*, *Informant* and *Best Friend*) [9]. For example, a vehicle assistance system can be classified as a *Guardian angel* if it can act highly autonomously and require low reciprocal engagement from the human.

Frameworks such as these provide a structured way of thinking about the nature of interactions between humans and intelligent systems. This can support the development of more targeted and effective design guidelines. Frameworks can also support designers directly by providing an overview of the possible types of interactions and essential design factors to consider.

3 SYSTEM DESIGN

We created three prototypes to represent each of the three human-AI interaction paradigms as described by Van Berkel et al. namely Intermittent, Continuous, and Proactive [23]. The purpose of these prototypes is to investigate factors of importance when considering different forms of interaction for a given product and to demonstrate the usefulness of the aforementioned paradigms in framing these factors. Each of the prototypes demonstrates a music application for personal use. They allow the user to enjoy music in different ways by interacting with an AI system through a browser interface. The prototypes are developed in JavaScript, and the source code is publicly available on GitHub[1].

3.1 Intermittent prototype

Intermittent interactions are initiated by the user providing an explicit request to the system, with the interaction ending when the system has responded to the user's request [23]. Following this definition, we designed a music application in which the user simply selects a song and then it is played by the system, as shown in Appendix A. The system recommends playlists based on the user's profile. Many popular music streaming services (e.g., Spotify, YouTube Music) work similarly. Due to practical constraints and the short-term interaction of our participants with our prototype, the system is designed to ask for the users' favourite music genres directly. The system then presents a collection of playlists to the user which matches their reported favourite genres. Finally, the user can browse and listen to songs within the presented playlists. This system is based on explicit logic, not any form of machine learning, and therefore is a simple case of symbolic AI [5].

[1] https://github.com/anders160196/Human-AI-Interactions

3.2 Continuous prototype

Continuous interactions require ongoing engagement from both the user and AI throughout the activity. The interaction is a process rather than a single task, in which the user and the system actively collaborate, exchange information, and coordinate actions [23]. Following this definition, we instead sought to encourage more sustained and evolving interaction between users and AI systems. The prototype uses t-distributed stochastic neighbour embedding (t-SNE) [24]. This technique can be used to map high-dimensional data points into a two or three-dimensional space for visualisation and interaction purposes. We used this technique to visualise the similarity between hundreds of songs as a visual 2D representation that the user can interact with. The similarity of the songs is calculated from a set of audio features, specifically 'acousticness', 'danceability', 'energy', 'instrumentalness', 'liveness', 'loudness', 'speechiness', 'tempo', and 'valence'. All these audio features were obtained for each song through the Spotify API [22]. When the user first interacts with this prototype, they are presented with a segment of the 2D representation, as shown in Appendix B. Through panning, the user can navigate the entire song collection. Hovering over any of the songs will start that specific song. This design enables the user to use the visual layout to explore the relationship between different songs and genres.

3.3 Proactive prototype

In contrast to Intermittent and Continuous interactions, Proactive interactions are initiated by the system rather than the user. The interaction will be triggered by a sensor or change in the internal system state, resulting in a system action [23]. Following this definition, we imagined a system that automatically plays music that matches the user's mood. For example, if the user is happy, the system should play happy music. As such, the system will be fully autonomous, does not require any input from the user, and supports a more passive form of music consumption. To implement such an interaction, the system records images of the user using a webcam. The images are sent to the Google Vision API [6], which runs a facial expression analysis and subsequently returns the results. If either joy, sorrow, anger, or surprise is detected in the users' facial expressions, the system will play a matching song, as shown in Appendix C.

4 METHOD

Three prototypes have been created to represent Intermittent, Continuous, and Proactive human-AI interaction. To evaluate the user experience across these interaction paradigms we conducted 12 qualitative in-person user interviews followed by a reflexive thematic analysis.

We recruited university students from a diverse set of educations, including Engineering, Medicine, Media Informatics, Sustainable Design and Techno-Anthropology. Six male and six female students between the ages of 19–27 (M = 24) participated in this study. Participants were not compensated for their participation.

Each interview was conducted as a 1-on-1 session, which lasted for approximately 30 minutes. First, the participants would answer general questions about their relationship with music and prior experiences with music systems. Next, they would test each of the three prototypes and provide comments such as initial reaction, perceived usefulness, usability and user experience. To mitigate any study fatigue and priming factors, we followed the Latin-square method and changed the order in which the prototypes were presented between every interview [11, p.177]. Finally, participants would compare all three prototypes and provide any further reflections in an open discussion.

Audio recordings, screen recordings and written notes were collected and stored from the interviews. The audio recordings were transcribed using machine transcription software, and the resulting transcripts were proofread and corrected by the first author to ensure accuracy. We then conducted a thematic analysis on the transcripts following the six-step process outlined by Braun and Clarke [2, p. 35-36]. The analysis involved familiarising ourselves with the data, highlighting and labelling relevant passages (coding), organising codes into themes, reviewing and refining themes, and writing the analytic narrative.

5 RESULTS

We next outline the primary themes emerging from our thematic analysis, which focus on the goals supported by different prototypes, the differences in perceived user control, and suggestions to enhance user control.

5.1 Diverging goals

All participants valued music as an important and meaningful activity. They listen to music in various situations (e.g., studying, exercising, cooking) and for multiple reasons (e.g., relaxation, entertainment). When comparing the three prototypes, they identified different supported goals. The Intermittent prototype was described as the most straightforward and similar to existing music apps. For that reason, it was also their preferred choice for listening to music regularly.

By comparison, the Continuous prototype required additional attention during the interaction. Therefore, this type of music system would not be suitable for users who are engaged in other activities, such as driving a car: *"I'm the kind of user who's going to be driving. So I cannot use a complex user interface, I just need one button"* (P1). However, participants believed that the Continuous prototype is great for actively exploring new music: *"I would say this is a quite fun way actually to try to discover new genres as well. [...] as you move the mouse, you might accidentally cross over one of the other ones and listen to that briefly. And that might catch your interest."* (P3) and *"It gives a fun sense of exploration of like, oh, there's more stuff over here. What kind of thing has it put together?"* (P2).

The functionality provided by the Proactive prototype was received with some scepticism. Participants argued that emotions are not always shown in your facial expression and that experiencing an emotion does not imply that any specific songs should be played. However, one participant suggested that the system could be useful when having friends over, and you want the music to be controlled not by a single person's mood, but the vibe of the whole group. Similarly, one other participant suggested using it as an AI DJ in a nightclub: *"What I would use something like this for is kind of an AI DJ, which you have cameras in a club and it looks at how the crowd responds to the music. And then what a DJ does is to look at the crowd* *and, okay, am I switching to a new song? What kind of song do I need?"* (P1).

In summary, the three prototypes were found to support users in achieving different goals. The Intermittent prototype was most useful for simply listening to music, the Continuous prototype for exploring new music, and the Proactive prototype for adapting to social settings.

5.2 Perceived user control

Our participants perceived clear differences in user control across the three prototypes. In the Intermittent prototype, the user can directly select which song to play from any of the recommended playlists (Figure 1). Participants found that this prototype provided the most direct control of the music.

For the Continuous prototype, a selection of about 20 songs is initially presented on the screen. To explore other songs, the user must drag the screen to reveal more songs, based on the layout that had been decided by the AI (Figure 1). In contrast to the Intermittent prototype, the Continuous prototype presents songs less systematically—inviting users to explore. Our participants largely experienced this as a lack of control, in part introduced by the perceived complexity of the user interface and unfamiliarity with its navigation style: *"As a user interface, I don't think this would work because it's quite complex."* (P1), *"There's too much going on and yeah, I think it's confusing that you need to drag it around to find some other types of music."* (P7).

In the Proactive prototype, the AI dictates which song to play based on the detected emotion of the user (Figure 1). P11 highlights some of the reasons for their scepticism towards this idea. First, users will try to control the music by altering their facial expressions, even though the system is designed to respond to their natural emotions: *"If I want some music that gets me in an angry mood, I have to already make myself angry."* (P11). Secondly, the emotion the user is feeling may not necessarily be the emotion they want the music to reflect: *"It's a neat idea, but it doesn't feel like it would support what I use music for. I don't really want music that supports the mood I'm already in. I want music to set a mood."* (P11). Generally, participants did not trust that this type of system would be able to meet their needs. Furthermore, the lack of user control could potentially result in negative experiences, as expressed by P2: *"If it just adjusts based on my mood, it's like, oh, you were listening to this song. I don't think it fits your mood anymore. That might feel frustrating if I've selected, I want to listen to this song."* (P2)

In short, each prototype evokes a different perception of user control. The Intermittent prototype offers the most direct control over music selection. Participants experienced a lack of control with both the Continuous and Proactive prototypes. In the Continuous prototype, this stemmed from the complex and unfamiliar interface. For the Proactive prototype, participants experienced a lack of control, because they did not trust a fully automated system to meet their needs.

5.3 Enhancing user control

Participants provided suggestions on how user control could be further enhanced in potential future iterations of the prototypes. For the Intermittent prototype, participants gave mixed responses on the performance of the AI in recommending playlists. One participant praises the accuracy of the AI recommendations: *"Damn, this is actually kind of good. This is actually quite accurate to what I listen to"* (P3), while another participant criticised the AI recommendations for being too presumptuous and uninspiring: *"So the recommendations are very pop music, very hip-hop music and very jazz music. Where I would think the interesting music is in between them, a mixture of jazz and hip-hop"* (P9). Recommended playlists need to not only match the users' general music preferences but also sometimes allow for the discovery of entirely new kinds of music. This dichotomy could explain why the Intermittent prototype received mixed responses, as it will only recommend playlists that fall within the users' preferred music genres. One participant suggested a new feature to the Intermittent prototype, which would extend the interaction between the user and AI system: *"Apple used to have a function where you could take five songs and then they would generate a playlist based on those five songs"* (P11). This feature would provide more control to the user, as they can decide not only which songs to play from recommended playlists, but also which input is used for generating new playlists.

The Continuous prototype was designed to promote active collaboration between the user and AI. However, one participant describes that the interaction does not feel like a collaboration: *"The second one you described as a collaboration, you do explore a bit, but it is like you are not taking, it doesn't feel like I'm taking any actions. I'm just like, oh, here's a map. Let's look at this and look at this"* (P2). While the prototype provides a 2D representation for the user to navigate, it is experienced as static. To make the interaction feel more collaborative, the participant suggested that the AI should be more dynamic and responsive to her actions: *"Stuff will pop up and it'll be like, oh, you made this decision. Let me suggest your next decision"* (P2). While the interaction in the Continuous prototype did not feel collaborative to some participants, feedback can play a crucial role in bridging this gap.

Finally, for the Proactive prototype, the AI dictated which song to play based on the detected emotion of the user. We found that participants would purposefully change their facial expressions in order to change the music. This behaviour conflicts with the proposed system design, which suggests a more indirect interaction. As such, it is not surprising that some participants would have issues with the level of control: *"I guess I would want more control over the music I listen to. I would want to say, okay, I want to listen to Happy music. Now using a camera for this is just an over-complicated interface in my head"* (P1). This is an example of misalignment between designer intent and user expectations [8, 17], a common pitfall when designing Proactive AI systems [15]. To avoid this misalignment issue, one participant suggested that some information should be hidden from the user: *"Then I don't think that there should be any interaction, in a sense, that I think the system should be observing me rather than me looking at it. [...] I mean, the same way I guess is here, but for me not to be able to see myself or the smileys. So that it does it without me knowing it, because the moment I see myself, I will be aware of myself"* (P4). By hiding information about what the AI has detected, we increase the autonomy of the system.

All of these participant suggestions demonstrate how the interaction paradigms can help frame possible design directions. Our user evaluations provided valuable insights into the nature of the three interaction paradigms, highlighting their potential benefits and limitations, and shedding light on the challenges and opportunities for future development.

6 DISCUSSION

We next interpret the results of our evaluation within the context of designing for human-AI interaction and the relevant literature within this space. Specifically, we discuss the implications of our findings for the design and development of AI systems and identify any limitations of the study.

6.1 Implications for design

Aligning with the users' goals and providing suitable forms of interaction are crucial considerations when designing for human-AI interaction. Our user evaluations reveal insights into the characteristics of different interaction paradigms, offering practical guidance for the design process.

The Intermittent paradigm is most suited for tasks with clear and known objectives. It allows users to maintain control over their actions, and user control can be further enhanced by providing adjustable parameters that empower users to tailor their experience. The Continuous paradigm, on the other hand, provides better support for tasks that involve exploration and discovery. It requires users to surrender some control, especially with complex interfaces. However, user control can be improved by designing intuitive interfaces with precise feedback mechanisms that enhance user understanding.

Finally, the Proactive paradigm should be reserved for situations where user control is not necessary or beneficial, as it necessitates users to entrust the AI with complete control. Mennicken's research highlights a nuanced perspective on this, showing that excessive proactivity may lead to user discomfort and fear of required interventions, underscoring the importance of balancing autonomy with user preferences for control [13, p.127].

Our findings suggest that the visibility of certain information can influence user behaviour and the perceived autonomy of the system. For instance, hiding what the AI has detected can increase the system's autonomy by preventing users from consciously altering their behaviour to manipulate the system's output. This aligns with the findings of Rezvani et al., who demonstrated that the level of information displayed on a user interface can significantly impact user performance and trust in the system [20]. Therefore, carefully considering the information displayed to the user is crucial in balancing system autonomy and user oversight in the Proactive paradigm.

6.2 Grappling with uncertainty and complexity

Yang et al. outlined two factors that contribute to the unique challenges in designing for human-AI interaction: *Capability uncertainty*, which refers to the unknowns surrounding the functionality and performance of AI systems, and *output complexity*, which pertains to the innumerable and quasi-random nature of the system's possible outputs [25].

Proactive interactions are initiated and controlled by the AI. Therefore, the Proactive paradigm may be more heavily impacted by capability uncertainty, due to this heavier responsibility. As exemplified in the Proactive prototype, in which the AI was tasked with predicting and fulfilling the user's needs, deciding when and which music to play. Participants expressed concerns regarding the Proactive prototype's capabilities, particularly its ability to understand complex human emotions and evolving music preferences—a feat they did not trust an AI to accomplish. Capability uncertainty is a challenge for designers and users, as they grapple with understanding what the AI can do in the real world [16]. In this light, Morrison et al. propose a more holistic view of capability in which capability emerges not solely from the technical system, but through the interaction between the user and AI [16].

The Continuous paradigm, characterised by ongoing engagement from both the user and AI, tends to amplify output complexity due to the dynamic and evolving nature of the interaction. This is evident in the Continuous prototype, where the arrangement of songs is not confined to a specific pattern but can be clustered in numerous ways. This flexibility allows any song to be placed anywhere within the visual representation created by the AI. The user can interact with this graphical representation, exploring any direction they choose. This level of variability and the broad range of possible outcomes indicate high output complexity.

6.3 Limitations & future work

In this study, we developed and evaluated three prototypes of a music application that represent three distinct interaction paradigms: Intermittent, Continuous, and Proactive [23]. While we carefully designed each prototype to capture the essence of each interaction paradigm, they provided only one representation. As such, other representations of these paradigms might have resulted in different participant perceptions. Additionally, we invite the exploration of interaction paradigms beyond music applications. This would involve investigating how these paradigms can be adapted to meet the unique user needs and complexities found in various contexts, such as healthcare or education, to design more intuitive and effective human-AI interactions.

Further, the test setup may not fully reflect natural user behaviour. This particularly impacted the Proactive prototype, as participants often acted out certain facial expressions to control the music. This goes against the main purpose of Proactive interaction, which is to minimise user effort. Future work should, therefore, consider more ecologically valid evaluations.

It should be noted that the participant group consisted of individuals from a relatively narrow age range and were all current university students. A more diverse participant sample regarding age and background would provide a broader perspective and potentially uncover additional insights into the user experience across different demographics.

Finally, given the positive engagement with the Continuous prototype, where users enjoyed exploring music without fully understanding the AI functionality, future work should delve into how partial transparency affects user experience. This approach challenges the prevailing emphasis on explainability, suggesting that facilitating user agency and interactive exploration may be equally important in enhancing the user experience with AI applications.

7 CONCLUSION

In this paper, we explored the challenges of designing for human-AI interaction in a music application. By developing and evaluating three prototypes, we compared end-user perceptions across three distinct paradigms of human-AI interaction. We found that each prototype offered varying degrees of user control, allowing designers to support divergent user goals and needs for control. Our findings underscore the importance of aligning the interaction paradigm with the application's primary purpose. The presented results contribute to a deeper understanding of the complexities in designing AI systems and offer insights for developing more user-centred AI applications. As such, we hope our work inspires further research in this area and inspires designers to consider alternative interaction paradigms to support divergent user goals.

ACKNOWLEDGMENTS

This work is supported by the Innovation Fund Denmark (IFD) under File No 9069-00038B in the project Interactive Sound Zones for Better Living (ISOBEL) and the Carlsberg Foundation, grant CF21-0159.

REFERENCES

[1] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300233

[2] V. Braun and V. Clarke. 2021. *Thematic Analysis: A Practical Guide.* SAGE Publications, New Delhi, India. https://books.google.dk/books?id=mToqEAAAQBAJ

[3] IBM Corp. 2022. Everyday Ethics for Artificial Intelligence. Available at https://www.ibm.com/watson/assets/duo/pdf/everydayethics.pdf (2022/12/22).

[4] Microsoft Corporation. 2022. Microsoft Responsible AI Standard, v2. Available at https://blogs.microsoft.com/wp-content/uploads/prod/sites/5/2022/06/Microsoft-Responsible-AI-Standard-v2-General-Requirements-3.pdf (2022/12/22).

[5] Mariusz Flasiński. 2016. *Symbolic Artificial Intelligence.* Springer International Publishing, Cham, 15–22. https://doi.org/10.1007/978-3-319-40022-8_2

[6] Google. 2023. Cloud vision API: Detect faces. https://cloud.google.com/vision/docs/detecting-faces

[7] Google. NA. Responsible Development of AI. Available at https://ai.google/static/documents/responsible-development-of-ai.pdf (2022/12/22).

[8] Marc Hassenzahl. 2003. *The thing and I: Understanding the relationship between user and product.* Springer Netherlands, Dordrecht, 31–42.

[9] Silvana Hinsen, Peter Hofmann, Jan Jöhnk, and Nils Urbach. 2022. How Can Organizations Design Purposeful Human-AI Interactions: A Practical Perspective From Existing Use Cases and Interviews. In *Proceedings of the Annual Hawaii International Conference on System Sciences*. Hawaii International Conference on System Sciences, Kauai, USA. https://doi.org/10.24251/hicss.2022.024

[10] Eric Horvitz. 1999. Principles of Mixed-Initiative User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Pittsburgh, Pennsylvania, USA) *(CHI '99)*. Association for Computing Machinery, New York, NY, USA, 159–166. https://doi.org/10.1145/302979.303030

[11] I. Scott MacKenzie. 2013. *Human-Computer Interaction: An Empirical Research Perspective* (1st ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[12] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. 2021. A Survey on Bias and Fairness in Machine Learning. *Comput. Surveys* 54, 6 (July 2021), 1–35. https://doi.org/10.1145/3457607

[13] Sarah Mennicken, Oliver Zihler, Frida Juldaschewa, Veronika Molnar, David Aggeler, and Elaine May Huang. 2016. "It's like living with a friendly stranger": perceptions of personality traits in a smart home. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Heidelberg, Germany) *(UbiComp '16)*. Association for Computing Machinery, New York, NY, USA, 120–131. https://doi.org/10.1145/2971648.2971757

[14] Leila Methnani, Andrea Aler Tubella, Virginia Dignum, and Andreas Theodorou. 2021. Let Me Take Over: Variable Autonomy for Meaningful Human Control. *Frontiers in Artificial Intelligence* 4 (Sept. 2021), 1–10. https://doi.org/10.3389/frai.2021.737072

[15] Christian Meurisch, Cristina A. Mihale-Wilson, Adrian Hawlitschek, Florian Giger, Florian Müller, Oliver Hinz, and Max Mühlhäuser. 2020. Exploring User Expectations of Proactive AI Systems. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4, Article 146 (dec 2020), 22 pages. https://doi.org/10.1145/3432193

[16] Cecily Morrison, Edward Cutrell, Martin Grayson, Anja Thieme, Alex Taylor, Geert Roumen, Camilla Longden, Sebastian Tschiatschek, Rita Faia Marques, and Abigail Sellen. 2021. Social Sensemaking with AI: Designing an Open-Ended AI Experience with a Blind Child. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 396, 14 pages. https://doi.org/10.1145/3411764.3445290

[17] Donald Norman and Andrew Ortony. 2006. Designers and users: Two perspectives on emotion and design. In *Theories and Practice in Interaction Design.* CRC Press, Boca Raton.

[18] Ana Paiva, Iolanda Leite, Hana Boukricha, and Ipke Wachsmuth. 2017. Empathy in Virtual Agents and Robots. *ACM Transactions on Interactive Intelligent Systems* 7, 3 (Sept. 2017), 1–40. https://doi.org/10.1145/2912150

[19] Inioluwa Deborah Raji, Andrew Smart, Rebecca N. White, Margaret Mitchell, Timnit Gebru, Ben Hutchinson, Jamila Smith-Loud, Daniel Theron, and Parker Barnes. 2020. Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain) *(FAT* '20)*. Association for Computing Machinery, New York, NY, USA, 33–44. https://doi.org/10.1145/3351095.3372873

[20] Tara Rezvani, Katherine Driggs-Campbell, and Ruzena Bajcsy. 2017. Optimizing interaction between humans and autonomy via information constraints on interface design. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, Vol. 2018-. IEEE, Yokohama, Japan, 1–6.

[21] Reza Shokri and Vitaly Shmatikov. 2015. Privacy-Preserving Deep Learning. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security* (Denver, Colorado, USA) *(CCS '15)*. Association for Computing Machinery, New York, NY, USA, 1310–1321. https://doi.org/10.1145/2810103.2813687

[22] Spotify. 2023. Web API reference: Spotify for developers. https://developer.spotify.com/documentation/web-api/reference/#/operations/get-audio-features

[23] Niels van Berkel, Mikael B. Skov, and Jesper Kjeldskov. 2021. Human-AI Interaction: Intermittent, Continuous, and Proactive. *Interactions* 28, 6 (nov 2021), 67–71. https://doi.org/10.1145/3486941

[24] Laurens van der Maaten and Geoffrey Hinton. 2008. Viualizing data using t-SNE. *Journal of Machine Learning Research* 9 (11 2008), 2579–2605.

[25] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-Examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376301

A INTERMITTENT PROTOTYPE

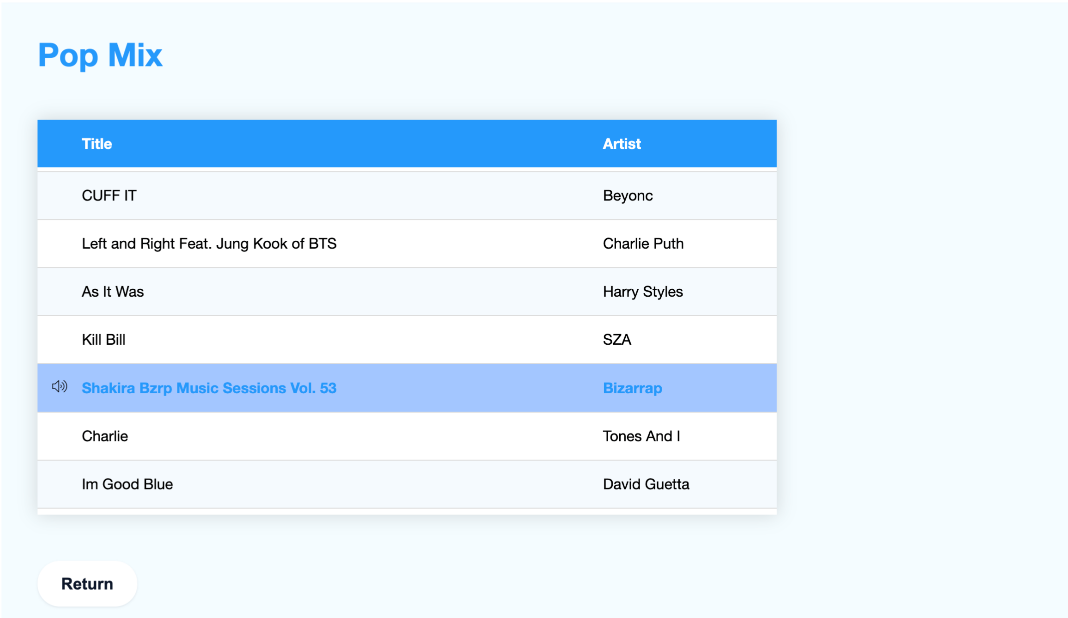

Figure 2: Intermittent prototype, which demonstrates an explicit user request followed by an AI response. The user chooses songs to play from a selection of playlists.

B CONTINUOUS PROTOTYPE

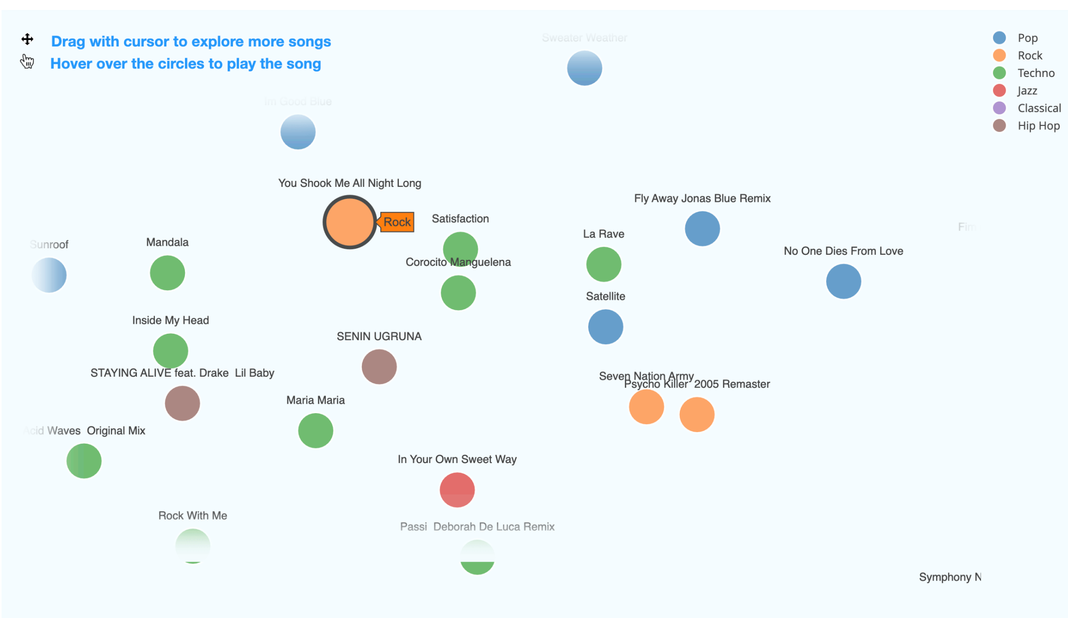

Figure 3: Continuous prototype, which demonstrates active collaboration between the user and AI. The AI facilitates user exploration through an interactive visual representation.

C PROACTIVE PROTOTYPE

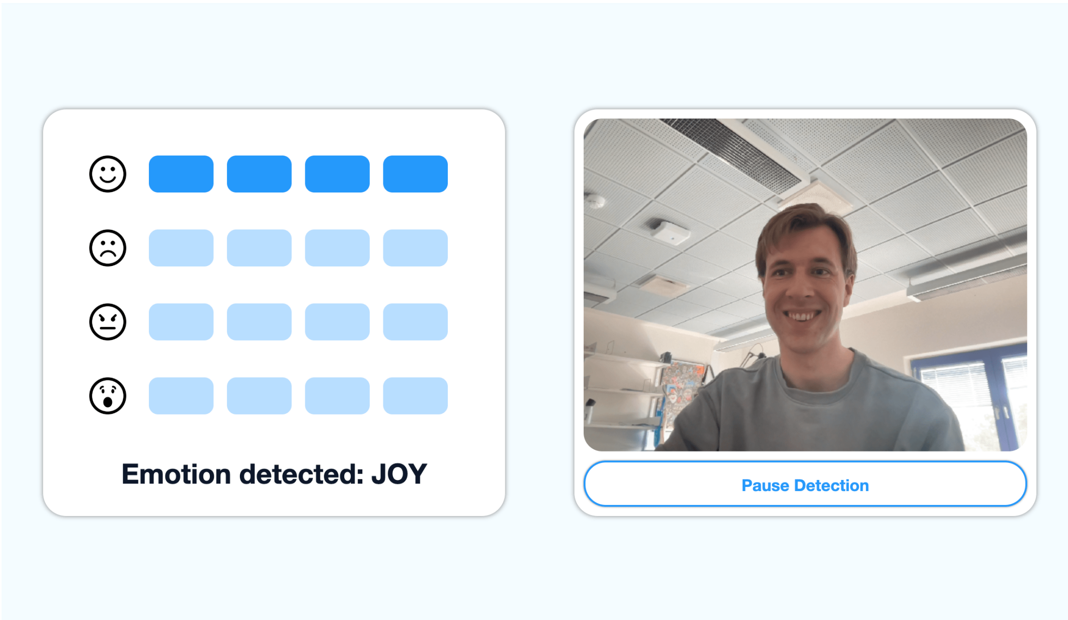

Figure 4: Proactive prototype, which demonstrates AI-initiated actions. The AI determines songs based on detected user emotions.