Aalborg Universitet

Human-AI Interaction: Intermittent, Continuous, or Proactive

van Berkel, Niels; Skov, Mikael B.; Kjeldskov, Jesper

Niels van Berkel, Mikael B. Skov, and Jesper Kjeldskov,
Aalborg University

Human-AI Interaction: Intermittent, Continuous, and Proactive

Insights

→ We have identified three types of human-AI interaction paradigms, which we refer to as intermittent, continuous, and proactive.

→ HCI researchers and practitioners need to systematically explore novel techniques to support end users in non-intermittent AI interaction scenarios.

With the rise in artificial intelligence (AI)–driven interactive systems, both academics and practitioners within human-computer interaction (HCI) have a growing focus on human-AI interaction. This has resulted in, for example, system-design guidelines and reflections on the differences and challenges when designing for AI-driven interaction as opposed to more-traditional applications [1]. We argue that the current work on human-AI interaction is defined primarily by a focus on what we refer to as *intermittent* interaction scenarios, in which there is a clear line between the human initiator of an interaction and an almost immediate system response. However, user interaction with AI systems does not necessarily follow this rigid interaction pattern. Inspired by Kristina Höök and Yang et al. [1,2], we define human-AI interaction as the completion of a user's task with the help of AI support, which may manifest itself in non-intermittent scenarios. By overlooking these other interaction paradigms, we neglect the opportunity to define and support alternative human-AI scenarios. In this article, we present and outline three types of human-AI interaction paradigms, which we refer to as *intermittent*, *continuous*, and *proactive*, highlighting a diverse set
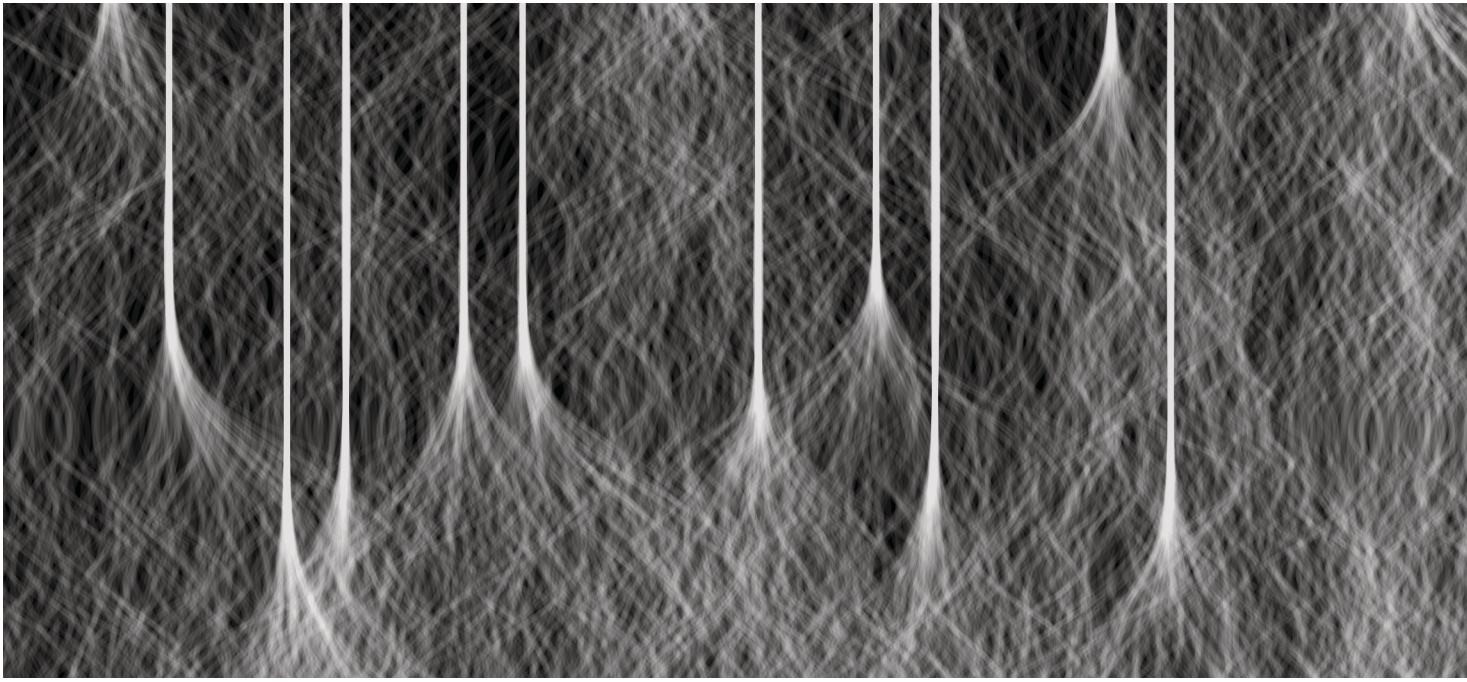
of interaction scenarios and pointing to a need for HCI considerations across different types of human-AI interaction. While a wide range of existing AI-powered systems operate continuously in the background of our lives (e.g., step counters, spam filters), these applications do not engage directly with their users. Here, we focus on AI applications that interact directly with their users.

No longer a faraway vision, AI increasingly augments our daily interaction with technology. After waking up, you might ask your phone, "Hey, Siri, how's the weather today?" to find out whether it will rain. During your commute, you open your music app and start an automatically curated playlist. When you're at work, your digital calendar identifies available meeting times to discuss an upcoming project with colleagues. Finally, as you wind down in the evening, you might ask your audio system to "play Enya in the living room." While these systems are technologically complex and offer different input modalities, the

interaction sequence between user and system is relatively straightforward: The user provides an explicit cue and input per the restraints of the system (e.g., weather request for current location; time slot and name of colleagues for a meeting). The system interprets the input, and subsequently formulates and presents a response in a predefined format. This response concludes the interaction until the user makes a new request.

In these examples, the interactions between the user and the system follow a turn-taking process, which is always initiated by the user (e.g., pressing a button, giving a speech command) and subsequently followed by a system response. We therefore categorize these as examples of intermittent human-AI interaction. Extending beyond the notion of AI support as a turn-taking process, many real-world applications of AI demand a continuous rather than an intermittent interaction sequence. For example, imagine an AI-powered driver training simulator—providing feedback and instructions as the user practices

their driving skills. The user would expect support from the system if an error were made during the simulated driving task. If the user would have to actively request AI support at every instance of the drive, the system would be unusable. Similarly, interrupting the user and preventing them from controlling the vehicle would hinder their ability to resolve challenging situations in a realistic manner. Many daily human tasks are continuous in nature, from playing a football match to making music. For an AI system to make a meaningful contribution in such activities, the system would have to continuously assess the stream of user input but only respond when necessary or relevant.

Furthermore, AI systems are increasingly proactive in assessing their context and autonomous in determining appropriate actions. Ubiquitous sensors and actuators allow these systems to respond to contextual changes without deliberate user input, mostly responding by themselves (for example, thermostats or lighting
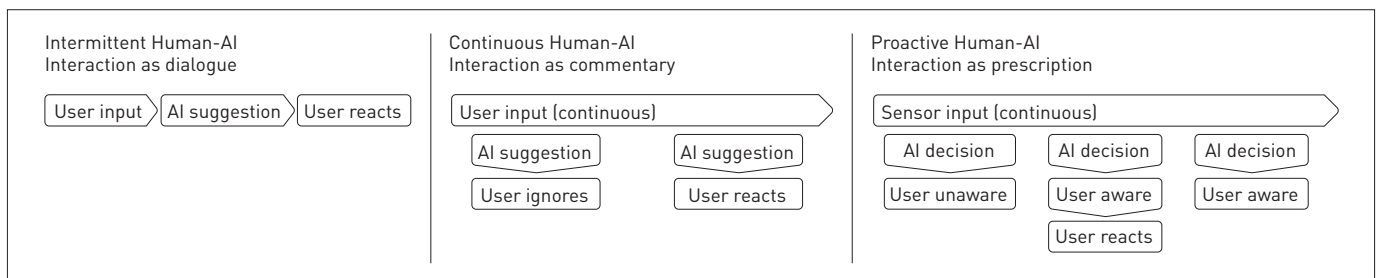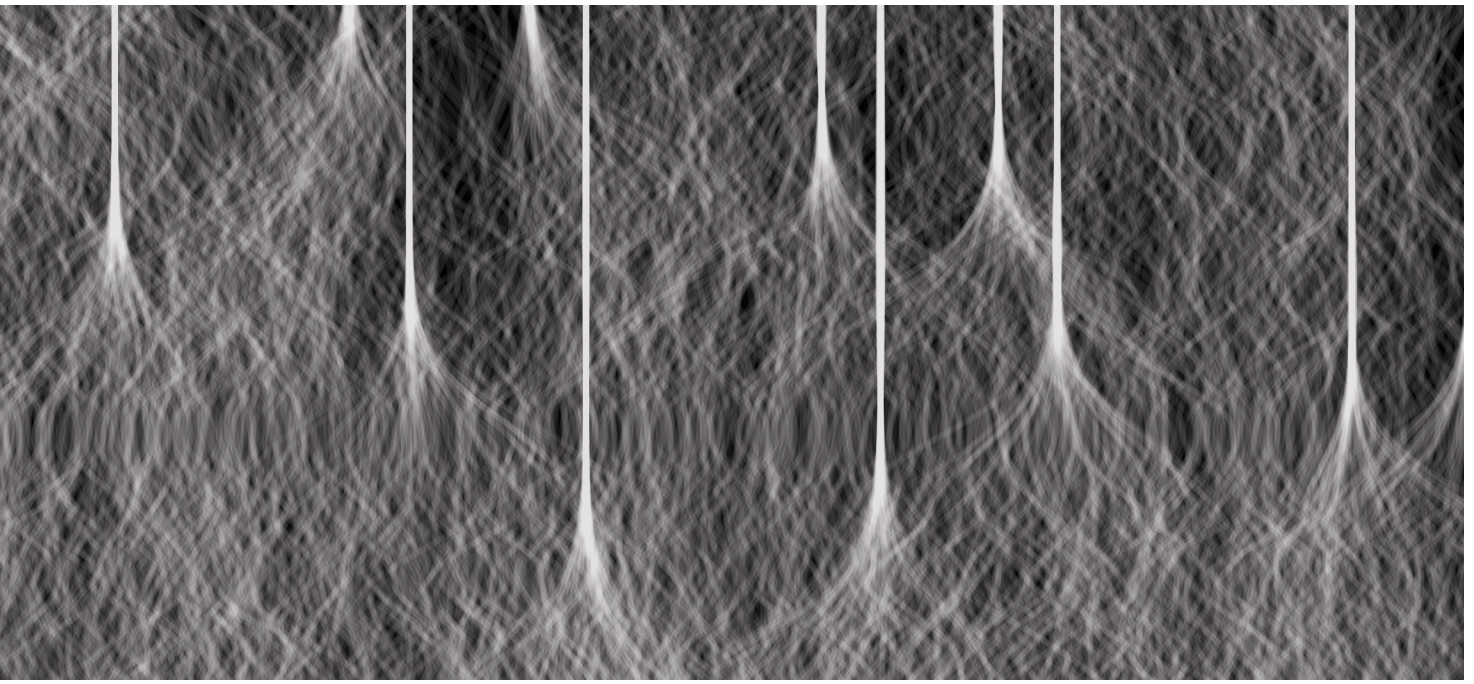
Figure 1. Visual representation of human-AI interaction across three paradigms of interaction.

systems that activate based on previously observed behavior or simply based on our presence in the context of our homes). These ubiquitous AI systems may, however, also require user input to complete an action or may seek this input from a user who is oblivious to the system's function, or presence, and therefore unaware about how to engage with it.

The current literature and study of human-AI interaction focuses primarily on turn-taking-based interaction, leaving other types of interaction underexposed (e.g., user input sustained over an extended period, user input is not the initiating factor). In this article, we describe the traditional perspective on HCI as a turn-taking process—as embodied in Donald A. Norman's action cycle [3]. We subsequently argue that the turn-taking paradigm prohibits a continuous interaction between users and AI systems in situations where the system is provided with an input stream rather than a singular user request. We provide examples of contemporary AI systems that follow this input paradigm and describe how the established perspective on user interaction breaks down in these scenarios. Looking ahead, we describe the notion of proactive human-AI interaction, in which AI systems do not necessarily await human input but rather initiate and drive user interaction. Given the stark differences between these three interaction paradigms, visualized in Figure 1, we argue that HCI researchers

and practitioners need to systematically explore novel techniques to support end users in non-intermittent AI interaction scenarios. We first present the three human-AI paradigms in more detail and describe how the current guidelines on designing for AI fall short on supporting human-AI. The article concludes with a call for future work toward continuous and proactive human-AI interaction to support future AI-based systems.

INTERMITTENT HUMAN-AI: INTERACTION AS DIALOGUE

Intermittent human-AI interaction describes systems that follow the notion of interaction between user and system as a turn-taking process. Norman introduced the *action cycle* as a description of the seven stages people go through when carrying out an action [3]. He describes these stages in the classic textbook *The Design of Everyday Things*, in which he presents the action cycle as applied to the scenario of reading a book:

Suppose I am sitting in my armchair, reading a book. It is dusk, and the light is getting dimmer and dimmer. My current activity is reading, but that goal is starting to fail because of the decreasing illumination. This realization triggers a new goal: get more light [3].

Following the identification of a goal (get more light, the first stage), the user goes through the execution phase in which they plan their action, specify an action sequence, and

perform the action. While an experienced user may carry out these stages subconsciously, they are still part of the action cycle. The execution phase is followed by an evaluation phase, in which the user perceives the updated state of the world, interprets this perception, and compares the outcome with the intended goal. If there is a mismatch between the outcome and the intended goal, the user may find themselves chasing a new goal (e.g., replace the light bulb). Norman's action cycle is based on the notion that it is the user who identifies a goal and initiates action.

The interaction between user and system follows a similar pattern when considering the current generation of intelligent digital assistants. As an example, with the goal of having silence for sleeping, the user plans to stop the music from playing and then specifies and performs the necessary action (e.g., saying, "Hey, Google, stop playing music"). The user then assesses whether the assistant has correctly carried out the desired action by perceiving and interpreting the new state compared with the previous one (has the music stopped playing?). As the interaction is initiated by the user and is followed by a response from the AI system, we describe this paradigm as *interaction as dialogue*. Another example of intermittent human-AI interaction is presented in Cai et al.'s pathology support tool [4]. Here, the system allows users to compare a

patient's tissue sample with historically similar samples. Throughout the query process, users can update the system's search criteria step by step, thereby steering the AI's focus on specific characteristics within the patient's sample when looking for historical similarities.

CONTINUOUS HUMAN-AI: INTERACTION AS COMMENTARY

Continuous human-AI interaction describes systems that "listen" to a stream of uninterrupted user input rather than individual instructions and can respond to this input throughout the duration of the interaction. Advances in processing speed and sensing capabilities have enabled new, continuous user-input capabilities. Rather than systems responding solely to explicit user input (e.g., button presses or audio triggers), continuous implicit user input is increasingly common. As an example, also highlighting that continuous interaction patterns are neither novel nor rare, real-time spelling and grammar correction in word processors are based on the user continuously inputting text, while the system continuously improves this (e.g., automatically indenting bullets in the text) or highlights elements that require further attention from the user (e.g., a spelling mistake).

The ability to continuously process a stream of user input over time enables new AI-support solutions. For example, van Berkel et al. studied AI support in the context of colonoscopy, in which medical experts are assisted in the detection of polyps by an image-recognition system [5]. Continuous user input (i.e., movement of the endoscope) provides continually updating video footage recorded from the patient's colon. This stream of input is sustained throughout the entire medical procedure; AI recommendations are suggested back to the user over time. It is essential to recognize that, in contrast to the intermittent human-AI

interaction concept, users are free to ignore the AI system's suggestions, and maintain their stream of input (Figure 1). As such, we portray this paradigm as *interaction as commentary*.

User input along a continuum requires a different approach to designing human-AI interaction. In these scenarios, the user is typically focused on a task and distracting them from this activity is undesired. Therefore, developers must find ways to integrate AI suggestions subtly into the user's ongoing task. Negative responses to, for example, Microsoft Office's Clippy point to the potential for user frustrations when interrupted while performing a highly focused task.

PROACTIVE HUMAN-AI: INTERACTION AS PRESCRIPTION

Proactive human-AI interaction describes systems that do not wait for user input but instead actively initiate and complete tasks based on, for example, sensor readings. Taking Norman's example of the reader of a book faced with increasingly dimmed surroundings, a relatively straightforward home automation system could increase the brightness of the surrounding lights when detecting both user presence and decreasing brightness levels. The execution phase, in which the user plans and executes their action, is thereby replaced by an automated system that aims to serve the user's goals.

Changes to the world (e.g., increasing light levels) are of course still perceived by the user, which subsequently leads to an interpretation by the user of the AI's action and an assessment of the desirability of the action taken. If the result is not deemed satisfactory, the user might initiate a full action cycle in which they adjust the light manually or alter the configuration of the home automation system. With the "bridge of execution" being removed from the interaction, the action cycle is reduced

to a "reaction cycle," in which the user responds to the behavior of proactive systems. While these proactive AI systems are intended to take work out of the user's hands and reduce their cognitive load by automating tasks, poor interaction between user and system can lead to extra effort being required by the user. We label proactive human-AI as *interaction as prescription*, as it is the system rather than the user that assesses the situation and decides to initiate an action.

LOOKING AHEAD

We presented three categories of human-AI interaction that are distinctively different in terms of user control and initiative: intermittent, continuous, and proactive. While intermittent human-AI interaction is an already widely established concept, the HCI community has not extensively engaged with the notions of continuous and proactive human-AI interaction. We summarize these types of human-AI interaction in Table 1, and present three arguments for why further attention from the community is required.

First, users have traditionally constructed mental models in which the user initiates the interaction and in which subsequent interactions follow a turn-taking process. These mental models do not apply to non-intermittent human-AI interaction scenarios. As stated by Stephen Payne, "users of machines are eager to form explanatory models and will readily go beyond available data to infer models that are consistent with their experiences" [6]. This desire to infer models on system operation will continue to manifest itself in non-intermittent forms of interaction. HCI researchers and practitioners therefore must assist users in constructing correct mental models to increase user understanding and prevent future errors. This is undoubtedly a challenging task in the context of systems that may change their behavior based on observed user behavior.

Second, proactive AI systems can reduce the friction we experience in our daily interactions with technology. However, this introduces new questions concerning user intentions and correction of system errors, as well as user consent and first-use experience. By focusing merely on intermittent interaction between users and their

The current literature and study of human-AI interaction focuses primarily on turn-taking-based interaction, leaving other types of interaction underexposed.

	Types of Human-AI Interaction		
	Intermittent	**Continuous**	**Proactive**
Initiator of interaction	User	User/AI	AI
Trigger for AI input	Explicit and predefined cue. For example, "Hey, Google."	Implicit, as part of constant input. For example, a video feed on which each frame is analyzed.	Change in system state matching a predefined condition. For example, movement captured through a sensor.
Resulting AI response	System response to user in predefined format. For example, "Temperature in Berlin is 18 degrees."	System response to user alongside the continuing user input. For example, a visual marker is introduced alongside a video feed.	System response to change in system state by changing the system. For example, turning on the lights and music when user comes home.
User response	User interprets the presented information. May be followed up by a new information request.	User decides to either ignore or react to the presented information, possibly adjusting the presented input stream.	User response not required; user might not be aware of AI intervention.

Table 1. Overview of three different types of human-AI interaction and their key differences.

devices, the community cannot address these critical questions that arise when deploying ubiquitous AI systems.

Third, the areas of fairness, accountability, and transparency have been raised as critical to the development of future AI systems—as shown by the recent establishment of the ACM Conference on Fairness, Accountability, and Transparency. How to embed these concepts in continuous or proactive human-AI interaction, however, is an open research question. In the context of continuous interaction, users may not want to receive transparent explanations on how the decision making came about, as it could distract from their primary task. In proactive AI scenarios, the user group is not tightly defined, with many (future) users of the system likely not present when decisions are made or explanations are provided. Furthermore, users' trust in systems that appear to adjust on a whim is likely to decline rapidly. Designing interactions around these settings will require further investigation from all areas of computer science.

CONCLUSION

Following the ever-increasing capabilities of computing devices, AI takes a growing role in supporting and driving human-system interactions. The increasing interest of both academia and industry in this space is an indicator of the expected benefits of AI integration for end users. We presented three distinct categories of human-AI interaction, highlighting how differences in initiation and control result in diverging user needs. These three paradigms of human-AI interaction can exist in parallel, and we do not necessarily believe that one of these paradigms will eventually disappear. Instead, it is likely that other types of human-AI interaction will manifest in the future. As continuous and proactive AI support systems are more likely to break existing mental models of end users, supporting designers in creating usable AI-driven systems becomes increasingly important. Therefore, we encourage future work in the areas of continuous and proactive human-AI interaction, ensuring the usability of the next generation of interactive systems.

ENDNOTES
1. Höök, K. Steps to take before intelligent user interfaces become real. *Interacting with Computers 12*, 4 (2000), 409–426; https://doi.org/10.1016/S0953-5438(99)00006-5
2. Yang, Q., Steinfeld, A., Rosé, C., and Zimmerman, J. Re-examining whether, why, and how human-AI interaction is uniquely difficult to design. *Proc. of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, 2020, 1–13; https://doi.org/10.1145/3313831.3376301
3. Norman, D.A. *The Psychology of Everyday Things*. Basic Books, 1988.
4. Cai, C.J. et al. Human-centered tools for coping with imperfect algorithms during medical decision-making. *Proc. of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, New York, 2019, 1–14; https://doi.org/10.1145/3290605.3300234
5. van Berkel, N., Ahmad, O.F., Stoyanov, D., Lovat, L., and Blandford, A. Designing visual markers for continuous artificial intelligence support: A colonoscopy case study. *ACM Trans. on Computing for Healthcare 2*, 1 (2020), 1–24; https://doi.org/10.1145/3422156
6. Payne, S.J. Mental models in human-computer interaction. In *The Human-Computer Interaction Handbook*. A. Sears and J.A. Jacko, eds. CRC Press, 2007, 63–76.

● **Niels van Berkel** is an assistant professor in Aalborg University's Human-Centered Computing group in the Department of Computer Science. His research interests lie in human-computer interaction, social computing, and ubiquitous computing.
→ nielsvanberkel@cs.aau.dk

● **Mikael B. Skov** is a professor in Aalborg University's Human-Centered Computing group in the Department of Computer Science. His research interests include pervasive and mobile computing, social computing, human-AI interaction, and usability engineering.
→ dubois@cs.aau.dk

● **Jesper Kjeldskov** is a professor in Aalborg University's Human-Centered Computing group in the Department of Computer Science. His research interests include indexical interaction design for context-aware mobile and pervasive computer systems and exploring new methods for studying the use of such technologies, both in the laboratory and in the field.
→ jesper@cs.aau.dk