AI Story Relay: A Collaborative Writing of Design Fiction to Investigate Artificial Intelligence Design Considerations

Hyeonjeong Im[*]
Soobin Jeon[*]
imhyeonjeong@kaist.ac.kr
soobinjeon@kaist.ac.kr
Industrial Design, KAIST
Daejeon, Republic of Korea

Haena Cho
haena.cho@kaist.ac.kr
Industrial Design, KAIST
Daejeon, Republic of Korea

Sungyong Shin
seneru@etri.re.kr
ETRI
Daejeon, Republic of Korea
Industrial Design, KAIST
Daejeon, Republic of Korea

Dasom Choi
dasomchoi@kaist.ac.kr
Industrial Design, KAIST
Daejeon, Republic of Korea

Hwajung Hong
hwajung@kaist.ac.kr
Industrial Design, KAIST
Daejeon, Republic of Korea

ABSTRACT

Artificial intelligence (AI) continuously evolves its level of complexity as it interacts with users. To address the unexpected outcomes of AI, developers desire to communicate with prospective users and reflect their opinions in the AI design pipeline. In this paper, we introduce the "AI story relay," a method by which the AI developer and users collaboratively construct a design fiction in a relay for AI design considerations. To explore the method's potential that encourages communication between AI developers and users, we conducted a case study of an AI-based coaching service. We expect AI story relay to contribute to fostering a dynamic exchange of views between developers and prospective users.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**.

KEYWORDS

Human-Computer Interaction(HCI), design method, participatory design, design fiction

[*]Both authors contributed equally to this research.

1 INTRODUCTION

Due to the lack of predictability in artificial intelligence (AI), AI-based technologies often cause undesired outcomes that have a considerable effect on society. For instance, a controversy arose around an AI chatbot named Tay that adopted users' racially and sexually offensive words [6]. To anticipate and prevent these unexpected consequences caused by user interaction with AI models, developers desire to understand the ways potential users and contexts shape AI's behavior [4]. Thus, the importance of active communication between AI developers and users in the AI design pipeline is growing [1].

Studies have provided opportunities to foster human-centered AI systems by allowing users to participate in the early development phase. For instance, Robertson et al. [5] improved the existing system that overlooked socioeconomic inequalities by involving stakeholders in the redesign of the algorithm that assigns students to public schools. Moreover, Lee et al. [3] redesigned an algorithm for food distribution to nonprofit organizations by making collective decisions with donors, recipients, and system operators. These approaches helped discover inherent user needs that are likely to be overlooked during the AI development process.

However, conveying users' opinions to developers in AI development practice is difficult. Without appropriate guidance, users unfamiliar with design approaches to image technology's future may have difficulty envisioning the potential risks of AI. Likewise, developers find it difficult to figure out which technical information they should provide users for successful practical discussions. Therefore, a new design method that allows non-technical users to imagine the application of AI technology easily is necessary to convey the users' perspectives effectively to developers.

In this study, we introduce the "AI story relay" method by which users can share tech-related perceptions and concerns with AI developers through collaborative design fiction writing. As a speculative design method, design fiction initiates discussions related to emerging technologies and facilitates users immersing themselves in the given context [2]. We further developed design fiction into a format in which developers and users write a story in relay to provoke a richer image that reflects various perspectives. As a case study, we performed this method to discuss a specific AI service, AI coaching

in sports. The case study showed that AI story relay helped the AI developer and users to express their underlying thoughts and envision a holistic view of the given service. We hope AI story relay provokes future fruitful discussions on the development of AI in practice.

2 METHOD

The AI story relay is a collaborative method where the AI developer and potential users participate together to explore and elicit AI design considerations. The method consists of three sessions: (1) setting the story background, (2) writing the story in the form of a relay race, and (3) conducting an interview. The study setup for each session is as follows.

2.1 Session 1: Setting the story background

The developer chooses a topic related to AI on which he or she wants to learn the users' perspectives. To enable users to fully understand the topic, the developer explains technology prior to writing the story. The explanation is restricted to facts to prevent any bias. Then, the developer and users co-create the story's settings, including the characters or situations. By building the story background together, the developer and users can immerse themselves in the story.

2.2 Session 2: Writing the story in relay

The developer and users write story fragments in a single relay round, and the relaying process iterates this round multiple times (Figure 1a). In the beginning, users write the story based on the initial descriptions of the AI. Users are encouraged to imagine ways a given technology could cause potential risks or social issues during this process. To support the story-writing process, we set three rules: (1) the participant can only write three to five sentences in one turn; (2) the last sentence should be left open-ended; (3) each turn has a time limit. The rules aimed to ease the burden on participants unfamiliar with story writing and to help them easily proceed with the unfinished story.

The developer proposes an updated version of the AI model at the end of each round reflecting on the users' reactions. We expected this AI update to provoke advanced discussions on the added features and to evaluate whether the developer fully understands the users' concerns. Multiple iterations would deepen these discussions and reflections.
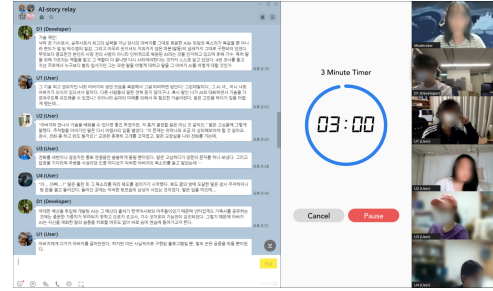
2.3 Session 3: Interview

After completing the story, an interview is conducted. The interview questions focus on the issues the users wanted to reflect on in the story, the ways they revealed their perceptions, whether their intentions were recognized and reflected by the developer, and whether the method influenced their perception of AI technology.

3 CASE STUDY

To explore the potential and challenges of the AI story relay method, we conducted a case study. We recruited five participants: one AI developer (D1) with two years of experience and four users (U1, U2, U3, and U4) with no AI-related technical background. The study was conducted through Zoom, a video conferencing platform. The
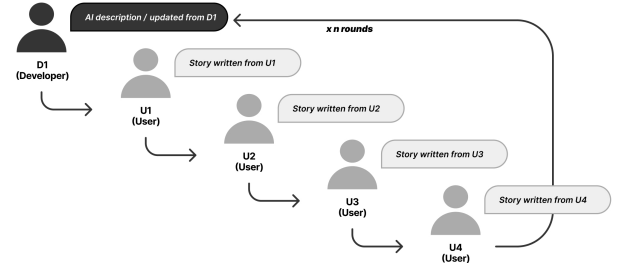
(a) Diagram of relaying process

(b) Case study sessions

Figure 1: Study settings for the AI story relay

developer and users simultaneously wrote stories in a group chat application (Figure 1b).

Through an interview beforehand with the AI developer, we selected as the case study topic an AI-related issue in which the developer was interested: an AI coaching service that aims to train junior jockeys. The story covered the relationship between an AI coach, created based on data from a deceased father who was a former jockey, his daughter, and his wife. The story began with the daughter entering a national jockey club and encountering an AI coach that resembles her deceased father.

The relay session involved five rounds, and the overall case study took two hours. The entire study was recorded, including video and audio, with consent. We analyzed the case study by conducting open coding with thematic analysis of the transcription of the recorded audio and story written by the participants.

4 FINDINGS AND DISCUSSION

4.1 Allowing the developer's holistic understanding of user perceptions

AI story relay encouraged the developer to consider broader aspects of technology, including not only technical requirements but also social relationships with prospective users. Our method provided the developer with an opportunity to simulate users' responses to the features served by AI. To investigate users' implicit thoughts from various viewpoints, the developer deepened the technological level of AI update every round and furthermore inquired about ethical issues: *"The researchers continuously collected data about the deceased and trained the AI to form a similar character. (D1)"* Moreover, we observed the developer using AI story relay in a progressive manner, seeking solutions for controversial issues such as *"What legislation on data collection and use will be needed in*

future AI applications? (D1)" and constantly heading toward a social consensus on the technology's position. Thus, the AI story relay facilitated the extension of the developer's perspective on AI from focusing merely on technology to considering social impacts in a holistic view.

Moreover, the story created through our method allowed the developer to understand the background of users' stances toward AI. Users' opinions were mostly implied in the behaviors or dialogues of characters in specific situations given by the developer. For instance, a user showed hostility toward the AI coach through the protagonist's (the daughter) words, *"This is illegal; the data lacks consent. They exploited my father's data. (U1)"* This response helped the developer not only grasp the user's negative stance but also comprehend the detailed reason for her opposition - the concern about *"appropriate data collection for AI training. (D1)"* As a result, the developer restricted the AI to *"collect data only from the words of the bereaved (D1)"* which sufficiently resolved the user's concern. As AI story relay allows users to imply their opinions in a social context, developers can wholly understand and put into consideration the reasons for such opinions.

4.2 Provoking users' critical thinking through exposure to various perspectives

AI story relay enabled users to establish critical viewpoints, regardless of their background in and knowledge of the technology. Users performed the method *"by adding a few extra individual thoughts to the ongoing story (U2),"* which allowed them to understand each other's perspectives and share personal opinions. With a dynamic exchange of views, some users completely changed the way they perceived the technology and empathized with their opposite stances. For instance, U4, who felt emotional when first meeting the AI, finally agreed to eliminate the AI permanently, saying, *"AI services created for the benefit of certain groups should be rejected. (U4)"* In addition, some users revealed concerns about possible exploitation of the given technology (e.g., *"such as current deep fake issues (U1)"* and *"What if personal data gets compromised in designing AI? (U3)"*). Although AI story relay mainly targets AI developers, such findings indicate that our method provokes critical thinking of users on the proposed technology.

5 CONCLUSION

This paper proposed a novel method to support communication between developers and users in AI development practice. AI story relay is a collaborative writing of design fiction that conveys users' perspectives of AI technology to developers. We discussed an AI-based coaching service from the case study and suggested the method's potential in our findings. AI story relay encouraged developers to understand and put into consideration the perceptions of users, and enhanced users' critical thinking about the social impacts of AI. However, the effectiveness of AI story relay highly depends on the users and their involvement; a solution to enhance consistency and stability should be sought. Therefore, future works should explore the directions of guidelines to assist developer interventions and provoke users' imagination. We expect our study to contribute to fostering a practical exchange of opinions between technology developers and prospective users.

REFERENCES

[1] Su Lin Blodgett, Solon Barocas, Hal Daumé III, and Hanna Wallach. 2020. Language (technology) is power: A critical survey of "bias" in nlp. *arXiv preprint arXiv:2005.14050* (2020). https://doi.org/10.48550/arXiv.2005.14050
[2] Derek Hales. 2013. Design fictions an introduction and provisional taxonomy. *Digital Creativity* (2013), 1–10. https://doi.org/10.1080/14626268.2013.769453
[3] Min Kyung Lee, Daniel Kusbit, Anson Kahng, Ji Tae Kim, Xinran Yuan, Allissa Chan, Daniel See, Ritesh Noothigattu, Siheon Lee, Alexandros Psomas, and Ariel D. Procaccia. 2019. WeBuildAI: Participatory Framework for Algorithmic Governance. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 181 (nov 2019), 35 pages. https://doi.org/10.1145/3359283
[4] Iyad Rahwan, Manuel Cebrian, Nick Obradovich, Josh Bongard, Jean-François Bonnefon, Cynthia Breazeal, Jacob W Crandall, Nicholas A Christakis, Iain D Couzin, Matthew O Jackson, et al. 2019. Machine behaviour. *Nature* 568, 7753 (2019), 477–486. https://doi.org/10.1038/s41586-019-1138-y
[5] Samantha Robertson, Tonya Nguyen, and Niloufar Salehi. 2021. Modeling Assumptions Clash with the Real World: Transparency, Equity, and Community Challenges for Student Assignment Algorithms. 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 589, 14 pages. https://doi.org/10.1145/3411764.3445748
[6] The Guardian 2016. *Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter.* Retrieved Sept 29, 2022 from https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter?CMP=twt_a-technology_b-gdntech