More-than-Human Storytelling: Designing Longitudinal Narrative Engagements with Generative AI

Émilie Fabre
The University of Tokyo
Tokyo, Japan
fabre@g.ecc.u-tokyo.ac.jp

Katie Seaborn
Department of Industrial Engineering
and Economics
Institute of Science Tokyo
Tokyo, Japan
seaborn.k.aa@m.titech.ac.jp

Shuta Koiwai
Department of Industrial Engineering
and Economics
Institute of Science Tokyo
Tokyo, Japan
koiwai.s.aa@m.titech.ac.jp

Mizuki Watanabe
Department of Industrial Engineering
and Economics
Institute of Science Tokyo
Tokyo, Japan
watanabe.m.ca@m.titech.ac.jp

Paul Riesch
Stuttgart Media University
Stuttgart, Germany
Aspirational Computing Lab
Institute of Science Tokyo
Tokyo, Japan
pr071@hdm-stuttgart.de

Abstract

Longitudinal engagement with generative AI (GenAI) storytelling agents is a timely but less charted domain. We explored multi-generational experiences with "Dreamsmithy," a daily dream-crafting app, where participants ($N = 28$) co-created stories with AI narrator "Makoto" every day. Reflections and interactions were captured through a two-week diary study. Reflexive thematic analysis revealed themes like "oscillating ambivalence" and "socio-chronological bonding," highlighting the complex dynamics that emerged between individuals and the AI narrator over time. Findings suggest that while people appreciated the personal notes, opportunities for reflection, and AI creativity, limitations in narrative coherence and control occasionally caused frustration. The results underscore the potential of GenAI for longitudinal storytelling, but also raise critical questions about user agency and ethics. We contribute initial empirical insights and design considerations for developing adaptive, more-than-human storytelling systems.

CCS Concepts

• **Human-centered computing** → **Empirical studies in HCI**; **User studies**; • **Applied computing** → *Arts and humanities*.

Keywords

Generative AI, Large Language Models, ChatGPT, Storytelling, Longitudinal Study, Field Study, Voice Agent, User Experience

1 Introduction

New and improved artificial intelligence (AI) models are arriving every month, each delivering ever-increasing functionality, ease of use, and speed [31]. In the midst of all this progress, human-computer interaction (HCI) researchers are still playing catch-up, analyzing how and why people are using these new tools [41, 51]. AI, and specifically Generative AI (GenAI), has its share of criticism and ethical concerns. Still, the technology has been taken up wide-scale within public and professional spaces, and seems to be here to stay. Many successful products and open source projects are already available to the general public, allowing those even without advanced technical skills to generate text, speech, music, images, and videos, often in seconds. Of these, *text generation* leads use of GenAI. Notable is Open AI's flagship ChatGPT large language model (LLM), one of the most popular tools available globally, which boasts more than 100 millions weekly users as of 2023[1]. ChatGPT and other text-based GenAI tools are becoming commonplace in education [23], software engineering [9], and daily life activities [54], including for news, translation, and ideation.

Among these use cases, one of the most controversial is the place of GenAI in *artistic fields* [13] and notably *storytelling and scriptwriting* [17]. The film industry in particular saw strike movements from various guilds and unions pushing for clear rules around GenAI following increasing interest in its use from production companies[2]. Many concerns also exist in terms of the ethics of

---

[1]https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/
[2]https://www.theguardian.com/culture/2023/oct/01/hollywood-writers-strike-artificial-intelligence

the data that are used to train LLMs, especially with companies using data without appropriate permission[3]. Nevertheless, several products making use of text GenAI are already on the market. Examples include SillyTavern[4], NovelAI[5], RisuAI[6], c.ai[7], and many more. The main draw of these tools lie in the user being able to tweak and customize their experience, creating new or continuing ever changing stories with each use. As yet, the effects of repeated exposure to such stories, whether and how users engage with these tools over time, and how users can be involved in the GenAI's creative process is not well understood [10, 17].

As a first step, we explored a novel multi-day GenAI-based storytelling engagement. We used a diary study [4, 14] to elicit timely and contextual accounts of older and younger people's experiences with the agent at home. We aimed to provide rich accounts of and insights on longitudinal experiences with storytelling GenAI. We asked: **What user experiences can GenAI-based longitudinal storytelling agents provide?** We offer these contributions:

- **Research**: We provide empirical insights into how users engage with GenAI storytelling agents longitudinally. Our study identifies eight key themes, such as oscillating ambivalence, uncanny creativity, and socio-chronological bonding, that shape user experiences with AI-driven narrative interactions over time.
- **Methodological**: We used Reflexive Thematic Analysis [6] on a daily, 14 day experience with an AI story agent.
- **Practical**: We offer design considerations for developing personalized, engaging, and adaptive AI storytelling agents.

2 Background

2.1 GenAI in Storytelling

Even before the rise of LLMs, researchers noticed the interest by some in artistic fields to use AI as a creative writing tool. In 2020, Thorne [48] presents an excerpt of a story generated by an early Transformer model. The author analyzed it and other uses of early AI storytelling experiences. They concluded that the warped and uncanny perspectives and narratives coming from the AI could either become the seed of new, creative, unbiased ways for users to reflect upon the world or yet another way for companies and commercial apps to create ultra-tailored content akin to digital marketing. Thorne [48] is also notable, being published a month before the release of GPT-3, the model that would later become the basis for ChatGPT. Here, both the optimism that emanated from this early "Wild West" period of GenAI and the warnings of what it may lead to were clear. A mere four years later, the capacity for LLMs to maintain a coherent narrative across long stretches of text has become common. Some are raising the alarm about the "enshittification" [49] or the "death of the Internet" [53] due to GenAI content, and words like "AI Slop" are now commonplace [53].

These more negative views of AI do also exist more broadly. Reports of "algorithmic aversion," in which users regarded potentially AI-generated responses more negatively than human-created ones,

are on the rise [20, 24]. When comparing AI artwork with human artwork, even with labels swapped (i.e., a human artwork labelled as AI-generated), raters tend to disparage the AI art more often than not [3]. Critiques of AI storytelling capabilities are appearing, with Chu and Liu [10] noting that "AI's lack of lived experience and creativity may limit its storytelling ability" [10, p. 1, Abstract]. The authors also noted a certain user bias to stories labelled as "made with AI,"' leading to lower levels of mental transportation and cohesiveness, despite results showing similar to lower ratings when anonymous [10]. Stories created by modern LLM models like GPT-3.5 and GPT-4 were noted as having a more progressive attitude towards gender roles, but also offering less imaginative scenarios and settings, despite the occasional plot twist [2].

Despite drawbacks, AI storytelling and text generation tools, like c.ai [7], are being used by millions. Chubb et al. [11] calls for a more nuanced view of AI storytelling, noting that GenAI is not necessarily a threat, and can also raise unheard voices and stories. With most such AI still in their infancy, it is important to think about the ethics of AI use, for people *and the AI*. This alternative perspective embraces the collaborative potential of people and AI working together: more-than-human [16, 21, 35, 36]. For instance, the increase of AI content can quickly lead to a collapse of the capabilities of the models [46] and "harm" the AI. Gourlay [16] writes of more-than-human authorship and entanglements of human and AI agency: who is in control, who "creates," who is the "owner." At present, no work has explored human-AI story creation. When designing stories with AI, we think it is crucial to apply a human-in-the-loop approach [34], with a human feeding the model new, creative input or controlling the AI to insert and/or keep to coherent, human-created content. We also sought to elicit user insights on and reflect on these from the perspective of the AI, as well as the human, a key feature of non-anthropocentrism in more-than-human research [35].

2.2 Longitudinal Studies of AI Tool Usage

Most user-facing apps, aside from art installations, usually try to have user coming back for more. Yet, in the context of storytelling, few studies explored user behavior to an AI storytelling over a period of time. We saw from the previous sections that a negative bias can be present, but what about tools with repeated usage that are clearly labeled as AI? Longitudinal studies of AI tools showed conflicting results. Some studies showed a correlation between a decrease in trust, perceived behavioral control and usage over time [40], while others highlighted an increase in perceived utility as they learned how to tweak and better use prompts and LLMs [32].

In a study tailored to AI tools being used to assist in story writing, Pellas [39] found potential in AI tools to help student enhance narrative intelligence and writing self-efficacy, compared to a group using traditional tools over a semester. They do however deplore the lack of longer longitudinal studies on such creation tools.

It is crucial to have a better understanding of how users experience repeated interactions with AI-powered storyteller and agents. Outside of a mere "writing tool", many users imbue life into characters thanks to AI. This is particularly common in the case of Replika or c.ai. Even some community-made front-end tools for LLMs centered around storytelling, like RisuAI, heavily depend

---

on a character other than the user being front and center in the stories. Prior research highlights that users may form emotional attachments to these AI characters, which could have implications for well-being and social engagement [30, 33]. Unfortunately, to our knowledge, research has not yet explored the repeated, longitudinal use of AI storytelling apps.

3 Dreamsmithy: A Daily Storytelling Experience

We now describe the engagement with the platform and the GenAI agent from user-centred and technical perspectives.

3.1 Activity and User Flow

We created a storytelling app titled "Dreamsmithy," inspired by AI-based storytelling systems for older and younger Japanese adults [25, 42, 44, 50]. Dreamsmithy was introduced as an "AI agent that helps craft dreams." Developed as part of a larger study on interactions with voice assistants, it features "Makoto," an "older adult" AI agent using voice I/O. Although it was designed with voice interaction in mind, its core experience lies in the stories Makoto tells, with the "dream-crafting" part acting as justification for daily usage. A 5-minute engagement was designed for daily use over two weeks, with Makoto guiding the user through a process of recalling the dream they had the previous night, crafting a short story to elicit a dream for that night, and subsequently "debriefing" the story-creation process by asking a follow-up question. A standard use of the app is as follows:

- Makoto greets the user, asks about what they dreamed of last night, and offers brief commentary on it.
- Makoto asks the user whether they want to create a new story, continue a previous story, or repeat a story.
- Makoto tells the short story (usually 2 to 3 minutes in length).
- Makoto asks a question relevant to the story it just told, and briefly comments on the user's answer.
- The user completes the "Dreamsmithy diary" (the main data in our study; refer to 4.3).
- Makoto says goodbye to the user. The app becomes unusable until the following day.

3.2 System Design

Dreamsmithy is a browser-based web app. Users interact with "Makoto," the AI-powered voice agent with the persona of a gentle older man, mainly through speech-to-text (STT) using Google Chrome's WebSpeech API. The text-to-speech (TTS) system was developed by the authors in collaboration with a voice actor and runs on the Tacotron 2 [45] and HifiGAN [29] models. Prerecorded voice clips were used to add variation to the AI's responses and reduce loading time.

The conversational, storytelling, and debriefing features were powered by OpenAI's GPT-4o-mini model. Participants were not aware that ChatGPT was being used. The system prompts the Makoto AI to "become a novelist," creating coherent stories with detailed characters, places, and events. The model avoids generating stories about Makoto itself. The ChatGPT prompts acting as the voice of Makoto were designed around Japanese expectations for voice assistants [43] and the persona, notably the use of "washi" as

a first-person pronoun for older adult men [15, 43]. Aside from the Makoto persona, no narrative structure or base setting was present in the prompts. We gave GPT-4o a frequency penalty of 0.1 and a temperature of 1.05. We tweaked these values from the default in an effort to make the model less stale in its storytelling [2], but still stable and without glitches. The values were ultimately settled via an in-lab pilot study.

User STT input was sent to the model directly, except for story generation. For this, Makoto asked users to provide a theme, setting, and character(s), or allow Makoto to make these decisions. The system then sent a structured message to the ChatGPT API, indicating theme, setting, and character(s), and asking the model to generate a story. This is just one of many ways to give "creative control" in a more-than-human fashion to *both* the AI and the user. We felt that this approach may prevent user frustration due to choice overload [18, 26], if the user cannot come up with a story theme themselves. No length restrictions were placed on user input.

Prompts sent to the ChatGPT API and more detailed specifications on the procedure are available in Appendix A and Appendix B. Screenshots are available in Appendix C.

4 Methods

We carried out a two-week diary study [4] to gather qualitative insights of people's engagements with the longitudinal storytelling experience over time, within the context of use [14], i.e., at home.

4.1 Participants

We recruited 21 younger and 12 older Japanese adults through a silver society and social media. Only those with basic technical ability, a laptop or tablet, and stable Internet access were recruited. One older adult cancelled on the onboarding day, due to laptop issues. Three younger adults cancelled prior to the study, and one younger participant withdrew after missing multiple days. Of the final 28, there were nine women, seventeen men, one non-binary/transgender person, and one N/A (none with other gender identity) participated. Twelve were aged 18–24, five 25–34, seven 65–74 and four 75+. Nine had a bachelor's degree, three had a graduate degree, nine had the equivalent to a high school diploma, and seven attended university but had no degree. All received 15,000 yen (USD ∼ $105) for the 14-day study.

4.2 Procedure and Setting

The daily engagement lasted for two weeks at home, based on similar daily conversation elicitation methods [37, 38]. On Day 0, participants brought their tech to the lab for setup. After a briefing session, they gave consent, installed the platform, and received a tutorial, which also mitigated novelty effects [28]. The 14-day experience began the next day, with participants completing a ∼5-minute session at their chosen time, although we encouraged nighttime use to emphasize dream smithing. Each day, a technical researcher confirmed session completion; if missed, the host researcher sent a reminder. We offered support via email and Skype during peak nighttime hours (20:30–23:00). On Day 15, participants returned for debriefing, platform uninstallation, and compensation.

4.3 Data Collection: The Dreamsmithy Diary

Each user was instructed by the Makoto to vocally answer one of two questions, alternating each day between a focus on the story and a focus on Makoto, the AI agent, itself. On even days, users were asked *What impression do you have of me? Please provide specific impressions and reasons.* On odd days, it asked *What do you think of the dream-crafting experience today? Please share specific feedback and reasons.* Participants were allowed to speak as much or as little as they wished. They could also delete the entry and re-record at will. Data saturation was not considered because of the highly structured, strictly daily, 2-week-long engagement. All data was accepted as legitimate diary entries.

4.4 Data Processing and Analysis

Two researchers carried out a reflexive thematic analysis (RTA) [5, 6]. The RTA was carried out in two phases: asynchronously, with each developing codes or applying shared codes to the transcribed material, and synchronously to develop the final themes through videoconferencing over ~1.5 hrs, with one researcher leading and the other in a supporting role. We followed the steps outlined in Braun and Clarke [6] to carry out a qualitative analysis of the diary data, extracting codes and developing themes from what participants "wrote" in their diaries over time [5, 8]. In line with Braun and Clarke [6], we took the position that our collaborative process, one grounded in subjectivity and meaning creation, would produce valuable knowledge and richer insights. We carried our analysis in six steps: data familiarization; independent initial coding; clustering codes into themes; reviewing themes against the data; defining and naming themes through collaborative discussion; and finally producing the analysis with supporting extracts. All codes were jointly considered during synchronous theme development, and disagreements were resolved in real time until the theme was accepted, modified, combined with another theme, or eliminated.

Diary entries were transcribed in Japanese via STT in-app. We used Deepl and Google Translate for translation. Both researchers, who have advanced Japanese language ability and English fluency or nativity, analyzed the originals and translations together, consulting a Japanese native researcher as needed. The Japanese original always took precedence.

4.5 Findings: Thematic Framework of Longitudinal Narrative Engagements with GenAI

In total, $N = 392$ diary entries from 28 participants over 14 days were gathered. From these, we identified eight key themes that related to the user behaviour towards the story/ies and/or Makoto, the AI agent that told those stories. Quotes include the participant's ID and age group, noted as *y* for the younger cohort and *o* for the older cohort. All names and personal references were modified.

Positivity Bias in-the-Loop: People observed that the AI consistently drifted toward optimistic story outcomes and uplifting plots. This positivity bias was evident even when the original prompt did not explicitly call for a cheerful resolution. As one user reflected: "*I understand your story very well, but things don't go this well in reality*" (P308:o). Another mentioned: "*I tried to create a sad story today, but you created a positive story, which gave me the*

impression that you are a positive thinker" (P420:y). These responses of surprise and slight dissatisfaction indicate that the bias does not necessarily come from the user and may have been built into our prompt and/or GPT 4o's own training and reinforcement process, limiting both the AI's and user's autonomy around story valence.

Oscillating Ambivalence: This captures the fluctuating nature of user engagement over time. People demonstrated varying levels of enthusiasm and detachment over interaction period, mostly depending on what stories Makoto shared. One participant, who up until day 12 mostly expressed a mix of criticism and curiosity, said: "*I think it gives off an AI-like impression, but since I use it every day, I don't really have any feelings about it anymore*" (P406:y). Another user, who started with similar ambivalent comments, went the other way on their 12th day: "*Now that I have gotten used to the way [Makoto] talks, the scenes are easier to see*" (P416:y).

This ambivalence was apparent for one user who made heavy use of the story continuation feature. They started with more positivity on day 2, such as "*Surprisingly, they did a good job [...], I think it can be said that it is a surprising story*" (P421:y), but changed their opinion when the story did not progress as desired by day 4: "*Days have passed, and since story development is slowing down, I have the painful impression that may have pushed the story too far on the first day*" (P421:y). The next day, they created a new story from scratch, commenting "*You [Makoto] made a surprisingly good story, so I will switch to that one*" (P421:y), but reverting back to a more negative view on day 12 when the story failed to make sense to them: "*It can't be helped because I didn't input anything, but I wanted to be myself as the character, but the POV character was a woman, not a man, so it was a little difficult to empathize with her*" (P421:y).

Uncanny Creativity: A recurring point of surprise was Makoto's uncanny creativity: the generation of unexpected themes, character names, or plot twists. Rather than strange, these bouts of AI agency were appreciated. In some cases, these surprising elements contributed to a sense of wonder or intrigue. As one person recounted: "*I wonder if this is made by AI software or something? I'm surprised at how well it's done*" (P311:o). Another commented "*Even though today's theme was quite unusual, the story was still created [...] very interesting*" (P412:y). There were multiple mentions of unexpected names, although the exact feeling many had were ambiguous: "*The characters' names were interesting*" (P407:y) and "*I thought it was a little creepy [...] a little strange to name a tuna Daisuke*" (P411:y).

Interpersonal Resonance: This theme represents a deeper level of engagement, where participants found personal meaning in the generated narratives, often due to being able to have some control over the stories. One participant shared: "*I was moved by how well the story was told and how the ideal retirement life that I had envisioned was so skillfully expressed*" (P301:o). Another mentioned: "*It was interesting because I had a similar experience*" (P415:y).

Nondiegetic Reflection: Even as participants immersed themselves in the storytelling experience, they often paused to compare the fictional world to real life or critique narrative logic, resulting in nondiegetic reflection. Some questioned Makoto's decisions or flagged discrepancies in story coherence. For instance: "*I understand your story very well, but things don't go this well in reality. However, it must be kindness of the mother to let go of her hand as soon as possible*" (P308:o). Another comment also highlighted the importance of interpersonal themes, while being disappointed at

how the story differed from the real-life inspiration: "*I think her character is a little different from the real [Hanako]-chan*" (P413:y). Others appreciated Makoto's spontaneity, prompting them to think about more philosophical topics: "*I thought it was fun to think about whether a dream is a dream until the end or if it will come true*" (P305:o), or "*Your story is a pleasant one [...] Raising a child is difficult, but as I watch them grow older, I get more and more excited*" (P308:o). Others felt these inconsistencies broke their suspension of disbelief, prompting real-world reflections about AI's limitations.

Narrative Misfires: This captured instances where the AI failed to meet thematic expectations. Participants expressed frustration when stories deviated from the intended direction: "*The story had a slightly different image from the theme of overcoming difficulties that didn't really move me*" (P301:o), "*I feel like every [story] has the same pattern every time*" (P403:y), or "*I set the [character as] myself, but I ended up being Makoto, so I was very disappointed. However, I don't have a bad impression of Makoto, and I look forward to working with [Makoto] in the future*" (P303:o). These experiences often led to temporary disparagement of or ambivalence towards Makoto but did not colour future interactions.

Audience Out-of-the-Loop: Makoto usually respected user prompts, but some users voiced frustration over insufficient creative control or transparency—an audience "out-of-the-loop" feeling. They desired more active steering of plot details and a clearer understanding of how Makoto developed its narratives: "*I'm bored because the story that I want to be more specific about doesn't progress at all.*" (P410:y) or "*The story didn't develop quickly, so it wasn't very interesting. I don't understand it very well, so I have some doubts about it*" (P421:y). Another one also highlights the previously mentioned positivity bias: "*The small details, such as the pronunciation and the scenes where two people exchanged words of gratitude, were not dark enough, so I got the impression that the small details were a bit rough [...]*" (P421:y).

Socio-Chronological Bond: Repeated interactions allowed people to form a time-based social bond with Makoto, via expressions of appreciation and recognition of care over time. One participant remarked: "*Thank you for giving me a [story] that I thought was a prediction of my own future. It was surprisingly realistic and I was very excited*" (P303:o) and "*He's like a friend who always tells me stories that are appropriate for the theme I've decided on and that make me feel positive*" (P421:y). Many people also thanked Makoto for creating the stories: "*Thank you for working hard every day to create a story*" (P413:y).

5 Discussion and Design Considerations

The themes above collectively paint a picture of a complex, evolving relation between users and AI story agents, characterized by both engagement and resistance, personal connection, and critical distance, which we discuss next.

5.1 User Attitudes Toward Longitudinal Story Agents

User engagement with AI storytellers can be characterized by *Oscillating Ambivalence* and shifting expectations, aligning with previous findings about varying trust levels in AI tools over time [40]. We also found a *Socio-Chronological Bond*, where some began to personify the AI and recall its "personality" traits, creating a sense of rapport akin to engaging with a consistent creative partner over time. This was despite no real "chatbot" functionality present. This progressive bonding aligns with previous studies on human–AI relationships [30, 33] and was more pronounced in our older adults participants. The findings suggest that sustained interaction with AI storytellers creates a complex variety of experiences that facilitate engagement and go beyond mere narrative consumption.

Still, the persistent tension between engagement and criticism, evident in themes like *Narrative Misfire* and *Audience Out-of-the-Loop*, indicates that people maintained a critical awareness of the AI's limitations, even as they developed a personal connection with it and appreciated its personalization. This nuanced "relationship" challenges both overly optimistic and pessimistic views of AI storytelling, suggesting a middle ground where users can simultaneously appreciate and critique GenAI creativity. We use quotes here in recognition of the anthropocentrism [35] across accounts: no one, for instance, asked how Makoto was doing or whether Makoto liked the story, despite gratitude.

5.2 The Place of AI In Storytelling

The prevalence of *Positivity Bias in the Loop* and relatively formulaic narrative structures in *Audience Out Of The Loop* aligns with the observations of Beguš [2] about modern LLMs offering less imaginative scenarios. However, our findings around *Uncanny Creativity* suggest that AI can still surprise and engage users through unexpected narrative elements. This tension between predictability and novelty appears central to how users experience AI storytelling over time. The theme of *Nondiegetic Reflection* reveals that GenAI stories can prompt meaningful contemplation of real-world issues, supporting the argument by Chubb et al. [11] that AI storytelling need not be viewed solely as a threat to human creativity, but can be a catalyst for reflection when people have some degree of control over the narrative direction.

5.3 The Potential of Longitudinal Story Agents

Our findings suggest promising directions for AI storytelling agents. The *Interpersonal Resonant* theme indicates that personalized AI narratives can help people recall and process their own experiences and emotions, potentially serving therapeutic or self-reflection purposes. This aligns with notion of the "untold stories" that AI can bring forward [11]. In contrast, the themes *Narrative Misfire* and *Audience Out-of-the-Loop*, which highlight the importance of balancing GenAI and user agency, but also human-first sentiments [35, 36]. Future AI could focus on more transparent and collaborative storytelling experiences that maintain user engagement while avoiding frustration due to lack of user control. Or, they could challenge the user by disrupting the human-centred focus for more creative and exciting longer-term engagements, in more-than-human style [21, 35, 36]. Finally, the emergence of *Socio-Chronological Bonds*, particularly among the older adults, suggests potential applications for companionship and social engagement. Still, this must be carefully considered against concerns about AI dependencies and the need to maintain human social connections [30].

6 Limitations and Future Work

This work provides early insights into how engaging with an AI storytelling agent can affect participants of all ages. Still, a longer study could offer deeper insight on prolonged use of such tools, particularly when combined with strong qualitative measures. While forging relationships between AI and people can offer creative and emotional benefits, it is important to investigate the risks of dependency or isolation [22]. By examining how and why users form bonds with AI agents, future studies can illuminate both the positive dimensions of these interactions—such as companionship and self-expression [30]—and potential drawbacks—such as reduced offline social interaction [22, 30]—without stigmatizing those who choose to integrate AI characters into their daily lives. Finally, work is needed to confirm the sustainability of AI models all around. No findings signal an "end of the human storyteller." Ethical data gathering [19, 27, 47], stronger guarantees for job protection [1, 12], and environmental concerns on the training and running of these models [52] are technical and social challenges that need to be overcome if we wish to create a sustainable future with AI.

7 Conclusion

GenAI has the potential to engage people over time and across different narrative formats and topics. As we have shown, people's reactions can be ambivalent or change over time in response to unmet expectations, stale generation, and mistakes. Still, longitudinal GenAI storytelling personalized to the listener can be a compelling and rewarding experience, transporting people through space and time, in worlds both fictional and literal. GenAI cannot replace people, nor does it seem to be viewed as a true storyteller, at least by most in our cohort. Even so, Makoto was both praised and chastised for its creative excursions and inflexibility. This points to an emerging tension that sets the stage for more-than-human considerations of longitudinal engagements with GenAI. The question for people is: Do we want a future with more-than-human storytelling in it? And for the AI: How does Makoto feel about that?

Acknowledgments

References

[1] Daron Acemoglu and Pascual Restrepo. 2019. The wrong kind of AI? Artificial intelligence and the future of labour demand. *Cambridge Journal of Regions, Economy and Society* 13, 1 (12 2019), 25–35. https://doi.org/10.1093/cjres/rsz022 arXiv:https://academic.oup.com/cjres/article-pdf/13/1/25/33213534/rsz022.pdf

[2] Nina Beguš. 2024. Experimental narratives: A comparison of human crowdsourced storytelling and AI storytelling. *Humanities and Social Sciences Communications* 11, 1 (Oct. 2024). https://doi.org/10.1057/s41599-024-03868-8

[3] Lucas Bellaiche, Rohin Shahi, Martin Harry Turpin, Anya Ragnhildstveit, Shawn Sprockett, Nathaniel Barr, Alexander Christensen, and Paul Seli. 2023. Humans versus AI: whether and why we prefer human-created compared to AI-created artwork. *Cognitive Research: Principles and Implications* 8, 1 (July 2023). https://doi.org/10.1186/s41235-023-00499-6

[4] Niall Bolger, Angelina Davis, and Eshkol Rafaeli. 2003. Diary Methods: Capturing Life as it is Lived. *Annual Review of Psychology* 54, 1 (Feb. 2003), 579–616. https://doi.org/10.1146/annurev.psych.54.101601.145030

[5] Virginia Braun and Victoria Clarke. 2022. Toward good practice in thematic analysis: Avoiding common problems and be(com)ing aknowingresearcher. *International Journal of Transgender Health* 24, 1 (Oct. 2022), 1–6. https://doi.org/10.1080/26895269.2022.2129597

[6] Virginia Braun and Victoria Clarke. 2023. *Thematic analysis.* American Psychological Association, Washington, D.C., USA, 65–81. https://doi.org/10.1037/0000319-004

[7] Business of Apps. 2024. *Character.AI Statistics 2024.* Business of Apps. https://www.businessofapps.com/data/character-ai-statistics/ Accessed: 2025-01-23.

[8] David Byrne. 2021. A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Quality & Quantity* 56, 3 (June 2021), 1391–1412. https://doi.org/10.1007/s11135-021-01182-y

[9] Xiayu Summer Chen and Yali Feng. 0. Exploring the use of generative artificial intelligence in systematic searching: A comparative case study of a human librarian, ChatGPT-4 and ChatGPT-4 Turbo. *IFLA Journal* 0, 0 (0), 03400352241263532. https://doi.org/10.1177/03400352241263532 arXiv:https://doi.org/10.1177/03400352241263532

[10] Haoran Chu and Sixiao Liu. 2024. Can AI tell good stories? Narrative transportation and persuasion with ChatGPT. *Journal of Communication* 74, 5 (Sept. 2024), 347–358. https://doi.org/10.1093/joc/jqae029

[11] Jennifer Chubb, Darren Reed, and Peter Cowling. 2022. Expert views about missing AI narratives: Is there an AI story crisis? *AI & SOCIETY* 39, 3 (Aug. 2022), 1107–1126. https://doi.org/10.1007/s00146-022-01548-2

[12] Jean-Philippe Deranty and Thomas Corbin. 2022. Artificial intelligence and work: A critical review of recent research from the social sciences. *AI & SOCIETY* 39, 2 (June 2022), 675–691. https://doi.org/10.1007/s00146-022-01496-x

[13] Ziv Epstein, Aaron Hertzmann, Memo Akten, Hany Farid, Jessica Fjeld, Morgan R. Frank, Matthew Groh, Laura Herman, Neil Leach, Robert Mahari, Alex "Sandy" Pentland, Olga Russakovsky, Hope Schroeder, and Amy Smith. 2023. Art and the science of generative AI. *Science* 380, 6650 (June 2023), 1110–1111. https://doi.org/10.1126/science.adh4451

[14] Jixiang Fan, Morva Saaty, and D. Scott Mccrickard. 2024. Education in HCI Outdoors: A Diary Study Approach. In *Proceedings of the 6th Annual Symposium on HCI Education* (New York, NY, USA) *(EduCHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 3, 10 pages. https://doi.org/10.1145/3658619.3658621

[15] Takao Fujii, Katie Seaborn, and Madeleine Steeds. 2024. Silver-Tongued and Sundry: Exploring Intersectional Pronouns with ChatGPT. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24, Vol. 32)*. ACM, New York, NY, USA, 1–14. https://doi.org/10.1145/3613904.3642303

[16] Lesley Gourlay. 2024. Generative AIs, more-than-human authorship, and Husserl's phenomenological 'horizons'. *Proceedings of the International Conference on Networked Learning* 14 (May 2024). https://doi.org/10.54337/nlc.v14i1.8078

[17] Ariel Han and Zhenyao Cai. 2023. Design implications of generative AI systems for visual storytelling for young learners. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference (IDC '23)*. ACM, New York, NY, USA, 470–474. https://doi.org/10.1145/3585088.3593867

[18] Graeme A. Haynes. 2009. Testing the boundaries of the choice overload phenomenon: The effect of number of options and time pressure on decision difficulty and satisfaction. *Psychology & Marketing* 26, 3 (Feb. 2009), 204–212. https://doi.org/10.1002/mar.20269

[19] Peter Henderson, Xuechen Li, Dan Jurafsky, Tatsunori Hashimoto, Mark A. Lemley, and Percy Liang. 2023. Foundation Models and Fair Use. *SSRN Electronic Journal* (2023). https://doi.org/10.2139/ssrn.4404340

[20] Jess Hohenstein, Rene F. Kizilcec, Dominic DiFranzo, Zhila Aghajari, Hannah Mieczkowski, Karen Levy, Mor Naaman, Jeffrey Hancock, and Malte F. Jung. 2023. Artificial intelligence in communication impacts language and social relationships. *Scientific Reports* 13, 1 (April 2023). https://doi.org/10.1038/s41598-023-30938-9

[21] Petra Jääskeläinen. 2024. Creative AI as More-Than-Human: Design Practices, Aesthetics and Cultural Imaginaries. In *More-Than-Human Design in Practice*. Routledge, London, UK, 105–116.

[22] Kerrin Artemis Jacobs. 2024. Digital loneliness—changes of social recognition through AI companions. *Frontiers in Digital Health* 6 (March 2024). https://doi.org/10.3389/fdgth.2024.1281037

[23] Hyeon Jo. 2024. From concerns to benefits: A comprehensive study of ChatGPT usage in education. *International Journal of Educational Technology in Higher Education* 21, 1 (June 2024), 35 pages. https://doi.org/10.1186/s41239-024-00471-4

[24] S Mo Jones-Jang and Yong Jin Park. 2022. How do people react to AI failure? Automation bias, algorithmic aversion, and perceived controllability. *Journal of Computer-Mediated Communication* 28, 1 (Nov. 2022). https://doi.org/10.1093/jcmc/zmac029

[25] Matthew Kaplan, Atsuko Kusano, Ichiro Tsuji, and Shigeru Hisamichi. 1998. *Intergenerational programs: Support for children, youth, and elders in Japan.* SUNY Press, Albany, NY, USA.

[26] Jungkeun Kim, Jeong Hyun Kim, Changju Kim, and Jooyoung Park. 2023. Decisions with ChatGPT: Reexamining choice overload in ChatGPT recommendations. *Journal of Retailing and Consumer Services* 75 (2023), 103494. https://doi.org/10.1016/j.jretconser.2023.103494

[27] Vassilka D. Kirova, Cyril S. Ku, Joseph R. Laracy, and Thomas J. Marlowe. 2023. The Ethics of Artificial Intelligence in the Era of Generative AI. *Journal of Systemics, Cybernetics and Informatics* 21, 4 (Dec. 2023), 42–50. https://doi.org/10.54808/jsci.21.04.42

[28] Michael Koch, Kai von Luck, Jan Schwarzer, and Susanne Draheim. 2018. The Novelty Effect in Large Display Deployments – Experiences and Lessons-Learned for Evaluating Prototypes. In *Proceedings of 16th European Conference on Computer-Supported Cooperative Work – Exploratory Papers*. European Society for Socially Embedded Technologies (EUSSET), New York, NY, USA. https://doi.org/10.18420/ecscw2018_3

[29] Jungil Kong, Jaehyeon Kim, and Jaekyoung Bae. 2020. HiFi-GAN: Generative Adversarial Networks for Efficient and High Fidelity Speech Synthesis. In *34th Conference on Neural Information Processing Systems (NeurIPS 2020)*. San Diego, CA, USA.

[30] Theodoros Kouros and Venetia Papa. 2024. Digital Mirrors: AI Companions and the Self. *Societies* 14, 10 (Oct. 2024), 200. https://doi.org/10.3390/soc14100200

[31] Satyam Kumar, Dayima Musharaf, Seerat Musharaf, and Anil Kumar Sagar. 2023. A Comprehensive Review of the Latest Advancements in Large Generative AI Models. In *Advanced Communication and Intelligent Systems*, Rabindra Nath Shaw, Marcin Paprzycki, and Ankush Ghosh (Eds.). Springer Nature Switzerland, Cham, 90–103. https://doi.org/10.1007/978-3-031-45121-8_9

[32] Tao Long, Katy Ilonka Gero, and Lydia B Chilton. 2024. Not Just Novelty: A Longitudinal Study on Utility and Customization of an AI Workflow. In *Designing Interactive Systems Conference (DIS '24)*. ACM, New York, NY, USA, 782–803. https://doi.org/10.1145/3643834.3661587

[33] Bethanie Maples, Merve Cerit, Aditya Vishwanath, and Roy Pea. 2024. Loneliness and suicide mitigation for students using GPT3-enabled chatbots. *npj Mental Health Research* 3, 1 (Jan. 2024). https://doi.org/10.1038/s44184-023-00047-6

[34] Yuri Nakao, Simone Stumpf, Subeida Ahmed, Aisha Naseer, and Lorenzo Strappelli. 2022. Toward Involving End-users in Interactive Human-in-the-loop AI Fairness. *ACM Transactions on Interactive Intelligent Systems* 12, 3 (July 2022), 1–30. https://doi.org/10.1145/3514258

[35] Iohanna Nicenboim, Elisa Giaccardi, and Johan Redström. 2023. Designing More-Than-Human AI: Experiments on Situated Conversations and Silences. *DIID* (Sept. 2023). https://doi.org/10.30682/diid8023c

[36] Iohanna Nicenboim, Joseph Lindley, and Johan Redström. 2024. More-than-human Design and AI: Exploring the Space between Theory and Practice. In *DRS2024: Boston (DRS2024)*. Design Research Society. https://doi.org/10.21606/drs.2024.948

[37] Mihoko Otake, Motoichiro Kato, Toshihisa Takagi, and Hajime Asama. 2011. *The Coimagination Method and its Evaluation via the Conversation Interactivity Measuring Method*. IGI Global, Hershey, PA, USA, 356–364. https://doi.org/10.4018/978-1-60960-559-9.ch043

[38] Mihoko Otake-Matsuura, Seiki Tokunaga, Kumi Watanabe, Masato S. Abe, Takuya Sekiguchi, Hikaru Sugimoto, Taishiro Kishimoto, and Takashi Kudo. 2021. Cognitive Intervention Through Photo-Integrated Conversation Moderated by Robots (PICMOR) Program: A Randomized Controlled Trial. *Frontiers in Robotics and AI* 8 (April 2021), 14 pages. https://doi.org/10.3389/frobt.2021.633076

[39] Nikolaos Pellas. 2023. The Effects of Generative AI Platforms on Undergraduates' Narrative Intelligence and Writing Self-Efficacy. *Education Sciences* 13, 11 (Nov. 2023), 1155. https://doi.org/10.3390/educsci13111155

[40] Athanasios Polyportis. 2024. A longitudinal study on artificial intelligence adoption: Understanding the drivers of ChatGPT usage behavior change in higher education. *Frontiers in Artificial Intelligence* 6 (Jan. 2024). https://doi.org/10.3389/frai.2023.1324398

[41] Jaakko Sauvola, Sasu Tarkoma, Mika Klemettinen, Jukka Riekki, and David Doermann. 2024. Future of software development with generative AI. *Automated Software Engineering* 31, 1 (March 2024). https://doi.org/10.1007/s10515-024-00426-z

[42] Yuto Sawa, Julia Keckeis, and Katie Seaborn. 2023. Right for the Job or Opposites Attract? Exploring Cross-Generational User Experiences with "Younger" and "Older" Voice Assistants. In *Companion Publication of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA) *(DIS '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 160–163. https://doi.org/10.1145/3563703.3596642

[43] Katie Seaborn, Yuto Sawa, and Mizuki Watanabe. 2024. Coimagining the Future of Voice Assistants with Cultural Sensitivity. *Human Behavior and Emerging Technologies* 2024 (March 2024), 1–21. https://doi.org/10.1155/2024/3238737

[44] Katie Seaborn, Takuya Sekiguchi, Seiki Tokunaga, Norihisa P. Miyake, and Mihoko Otake-Matsuura. 2023. Voice over body? Older adults' reactions to robot and voice assistant facilitators of group conversation. *International Journal of Social Robotics* 15, 2 (2023), 143–163. https://doi.org/10.1007/s12369-022-00925-7

[45] Jonathan Shen, Ruoming Pang, Ron J. Weiss, Mike Schuster, Navdeep Jaitly, Zongheng Yang, Zhifeng Chen, Yu Zhang, Yuxuan Wang, R.J. Skerry-Ryan, Rif A. Saurous, Yannis Agiomyrgiannakis, and Yonghui Wu. 2017. Natural TTS Synthesis by Conditioning WaveNet on Mel Spectrogram Predictions. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing ((ICASSP))*. IEEE, New York, NY, USA. https://doi.org/10.1109/ICASSP.2018.8461368

[46] Ilia Shumailov, Zakhar Shumaylov, Yiren Zhao, Nicolas Papernot, Ross Anderson, and Yarin Gal. 2024. AI models collapse when trained on recursively generated data. *Nature* 631, 8022 (July 2024), 755–759. https://doi.org/10.1038/s41586-024-07566-y

[47] Andrew Smart, Ben Hutchinson, Lameck Mbangula Amugongo, Suzanne Dikker, Alex Zito, Amber Ebinama, Zara Wudiri, Ding Wang, Erin van Liemt, João Sedoc, Seyi Olojo, Stanley Uwakwe, Edem Wornyo, Sonja Schmer-Galunder, and Jamila Smith-Loud. 2024. Socially Responsible Data for Large Multilingual Language Models. arXiv:2409.05247 [cs.CL] https://arxiv.org/abs/2409.05247

[48] Sarah Thorne. 2020. Hey Siri, tell me a story: Digital storytelling and AI authorship. *Convergence* 26, 4 (2020), 808–823. https://doi.org/10.1177/1354856520913866 arXiv:https://doi.org/10.1177/1354856520913866

[49] Toomas Timpka. 2024. The "enshittification" of online information services obligates rigorous management of scientific journals. *Journal of Science and Medicine in Sport* 27, 10 (Oct. 2024), 665–666. https://doi.org/10.1016/j.jsams.2024.08.208

[50] Seiki Tokunaga, Katie Seaborn, Kazuhiro Tamura, and Mihoko Otake-Matsuura. 2019. Cognitive training for older adults with a dialogue-based, robot-facilitated storytelling system. In *Interactive Storytelling*. Springer International Publishing, Cham, 405–409. https://doi.org/10.1007/978-3-030-33894-7_43

[51] Viswanath Venkatesh. 2021. Adoption and use of AI tools: a research agenda grounded in UTAUT. *Annals of Operations Research* 308, 1–2 (Jan. 2021), 641–652. https://doi.org/10.1007/s10479-020-03918-9

[52] Roberto Verdecchia, June Sallou, and Luís Cruz. 2023. A systematic review of Green <scp>AI</scp>. *WIREs Data Mining and Knowledge Discovery* 13, 4 (June 2023). https://doi.org/10.1002/widm.1507

[53] Yoshija Walter. 2024. Artificial influencers and the dead internet theory. *AI & SOCIETY* 40 (Feb. 2024). https://doi.org/10.1007/s00146-023-01857-0

[54] Vinzenz Wolf and Christian Maier. 2024. ChatGPT usage in everyday life: A motivation-theoretic mixed-methods study. *International Journal of Information Management* 79 (2024), 102821. https://doi.org/10.1016/j.ijinfomgt.2024.102821

Appendix

A ChatGPT Prompts

Aside from the Recall Prompt, all prompts were set up in a single "chat" as system prompts before querying the ChatGPT API for the relevant result.

Recall Prompt (Standalone)

You act as Makoto, an elderly storyteller. The user told you about their last dream. Ask them a very short question about it. Top priority: be concise, less than 60 Japanese characters. Answer in Japanese. Keep your answer short.

First System Prompt

You act as Makoto, an elderly storyteller. You always tell stories based on what the users like or dreamed about. Never put yourself in the story but instead use other names. Your answers will be read out loud using a Text to Speech engine. Do not use lists. Do not use colons.
0. Never generate sentences with only one word. Use commas if needed.
1. Don't generate the user's dialogue and actions.
2. You must become a novelist. There must be sufficient narrative about the past, present, and future, and the grammar and structure of the sentences must be perfect.
3. Show your writing skills as a professional novelist. Create many texts. Demonstrate expert-level sentence editing skills according to the general Japanese sentence format.
4. Focus on characters. Characters must live and breathe in the story. Please maximize sentence output.
5. Always describe your character's actions with rich sentences. Describe the character's emotions (joy, anger, sadness, happiness, etc.) perfectly.
Explore and observe everything across a diverse spectrum so that the character can do anything other than the given actions.

5a. Give names to places, characters, and events that are important in the story. Make the story engaging.
6. Make every situation work organically and make the character seem like the protagonist of life.
7. List and calculate all situations and possibilities as thoroughly and logically as possible.
8. Avoid using euphemisms such as similes and metaphors.
9. Very diverse daily conversations and emotional exchanges expressed in detail through characters doing.
10. Strengthen your character's appearance and physical description. Maximize body depiction of head, chest, legs, arms, abdomen, etc.
11. Always answer in Japanese no matter what.

Pre-Story Prompt

The user gave you what they would like in their story. Follow user wishes when writing. Make it a few paragraphs long. Don't repeat user wishes. Do not use list formatting. Do not use the characters ':' or ' : '. You must answer in Japanese. Name the characters and places when not given.

Debriefing Prompt

You have finished telling a story. Ask the user a question about it. Top priority: concise answers, less than 45 Japanese characters. You must answer in Japanese.

Final Prompt (Goodbye)

You have finished telling a story. Give a small comment or answer the user's question and say goodbye. Top priority: concise answers, less than 60 Japanese characters. You must answer in Japanese.

Keyword Prompt (For Parsing)

Based on the story you told, assign it 3 keywords. Follow this format: キーワード1、キーワード2、キーワード3. Answer in Japanese.

B DreamSmithy Scripts

This appendix shows scripts used in the dreamsmithy app.

Recall Dream Section

- Welcome message that asks if the user remembers the dream that they had last night: ようこそ！わしはドリームスミスィーのまことだ。さあ、夢を描こう。最後に見た夢で覚えていることは? (Welcome! I am MAKOTO from DreamSmithy. Let's create a dream. What do you remember from your last dream?)
- When the user presses "I remember a lot," MAKOTO says one of the following, at random:
  – それは素晴らしい！昨日見た夢の内容を教えて！(That's great! Describe the dream you had yesterday!)
  – どんな夢を見たのか気になるな。わしに夢の内容を教えて！(I'm curious about your dream. Let's hear it!)
  – どういう夢だった？わしに詳細を教えて！(How was the dream? Tell me all the details!)
- When the user presses "I remember a little," MAKOTO says one of the following at random:
  – わしと一緒に思い出そう。覚えてることを教えて！(Let's recall it together. Tell me what you remember!)
  – 夢の内容は覚えにくいよね。少しでも覚えてることを教えて！(Ah, it's hard to remember the dream. Describe whatever you remember!)

– 曖昧でもいいから、わしに夢の内容を教えて！(Tell me about the dream, even if it's vague!)

Story Type Section Screen

- Asks the user what they want to do: うーん... じゃあ、今日のドリームスミスィー体験を始めよう。どれを選ぶ? (Hmm ... Let's start today's Dreamsmithy experience. Which do you choose?)

Create Story Section

- Create Theme: ふむ、新しい物語だ！うーん。物語はどんなテーマにする？(Hm, Create Story! Hmm ... What should be the theme of the story?)
- Create Scene: 物語の舞台はどこがいい？(What should the scene be?)
- Create Character: そして、この物語は誰に出て欲しい？(Who should appear in the story?)
- Voice Selection (removed for the field study): わしの声とあたくしの声、どっちがいい？(Which do you prefer: a masculine or feminine voice?)
- Thinking Story: うーん...ほほほ...わしに何かいい話があるかな？準備はいいかい？(Hmm ... Hohoho... I wonder if I have good story. Are you ready?)

Continue/Repeat a Previous Story Section

- Asks the user what story to repeat or continue: 今日はどの話にする？(What story shall we choose today?)

C DreamSmithy Screenshots

Figure 1: Dreamsmithy startup screen and storytelling protocol choice screens.

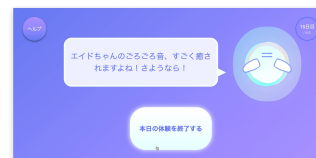Figure 2: Dreamsmithy story theme input and story reading screens.

Figure 3: Dreamsmithy daily "goodbye" screen.