Dynamic Participatory Game Design with Local AI: From Interviews to Trauma-Aware Interactive Narratives

Kseniia Harshina[1], Tom Tucek[1] and Mathias Lux[1]

[1]ITEC, University of Klagenfurt, Klagenfurt, Austria

Abstract

We present a work-in-progress participatory storytelling pipeline for sensitive narrative elicitation and replay in contexts of forced migration and displacement. The pipeline uses a locally hosted large language model (LLM) as a neutral chatbot interviewer. The system supports self-paced narration without cloud processing, prioritizing privacy, data sovereignty, and participant control. Interview transcripts are transformed into a structured scene representation (extracted fields and dialogue prompts), which is then replayed through a lightweight prototype interface as an initial step toward interactive memory-based experiences. We report a small formative expert evaluation (n=2) focusing on perceived comfort, emotional safety, and usability. Participants described the interviewer as low-pressure and reflective, while highlighting limitations such as weak acknowledgement of long answers and occasional "forced turns." We discuss design implications for narrative extraction, turn-taking, and staged evaluation in sensitive contexts, and outline next steps for community-informed studies with participants who have lived experience of forced migration or displacement.

Keywords

Large language models, Trauma-aware game design, Displacement, Narrative extraction

1. Introduction

Stories of forced migration and displacement are frequently under-documented, fragmented, or lost, particularly when people lack safe opportunities to record memories or when formal archival and interview infrastructures are inaccessible. At the same time, recounting personal experiences can be emotionally demanding, and conventional interviewing can introduce social pressure and power asymmetries.

This short paper presents a work-in-progress system that supports self-paced storytelling through a locally hosted, large language model (LLM)-powered chatbot interviewer. The interviewer is designed to be neutral and non-judgmental, and to operate without cloud infrastructure in order to prioritize privacy and data sovereignty. Interview logs are then transformed into a structured scene representation (extracted fields and dialogue prompts), which can be replayed through a prototype interface as an initial step toward interactive memory-based experiences.

Our approach sits at the intersection of ethical AI, trauma-aware design, and participatory storytelling [1]. Rather than framing the system as a replacement for human interviewing, we explore how local AI can support participant control and staged research workflows in sensitive contexts.

This work makes the following contributions:

- A locally hosted AI-based automated interviewer for sensitive memory elicitation, emphasizing privacy and participant control;
- An end-to-end pipeline from interview logs to a structured scene representation (schema and dialogue prompts) and a prototype replay interface;
- A formative expert evaluation (n=2) reporting perceived comfort, emotional safety, and usability issues to guide iterative refinement.

The remainder of the paper describes the system design, outlines the narrative extraction and replay components, reports formative evaluation insights, and discusses limitations and future work.

In: R Campos, A. Jorge, A. Jatowt, S. Bhatia, M. Litvak (eds.): Proceedings of the Text2Story 26 Workshop, Delft (The Netherlands), 29-March-2026

kseniia.harshina@aau.at (K. Harshina); tom.tucek@aau.at (T. Tucek); mathias.lux@aau.at (M. Lux)

2. Related Work

Automated interviewing and ethical AI. LLMs are typically optimized for answering questions rather than conducting interviews. Using an LLM as an interviewer inverts this paradigm. Prior work in conversational agents has explored empathy simulation and co-creative storytelling (e.g., Serbanescu [2]). More recently, research has also begun to examine LLMs as adaptive interviewers and conversational survey tools [3], as well as interview-based multi-turn evaluation frameworks for LLMs themselves [4]. Our work differs from these directions by focusing on locally deployed, trauma-aware narrative elicitation in a sensitive context, and on transforming interview logs into replayable interactive scene representations rather than only collecting survey-like responses or evaluating model behavior. We also draw on recent work in procedural narrative generation using generative AI [5, 6], while distinguishing our approach through its grounding in participant-authored experiences rather than AI-authored content.

Preserving personal and displaced histories. Historical accounts often rely on documents, institutions, or archives, yet stories of displacement are frequently absent from these records. Games such as *Bury Me, My Love* [7] and *Path Out* [8] explore migration experiences through authored narratives, but not as participant-authored, co-created experiences. This research also responds to critiques of stereotypical portrayals of suffering in mainstream media [9].

Trauma-aware and participatory game design. Trauma-aware design emphasizes emotional safety, agency, and care. Replayability can support reflection and narrative reframing, particularly when users revisit their own accounts. Participatory game design, as advanced by Flanagan [10], emphasizes co-authorship and inclusion of marginalized voices. Recent work has shown how autobiographical game making can support recovery and empowerment [11]. Our system combines these directions through an interview-to-scene pipeline that supports revisiting and iterating on participant-authored memories.

3. Methodology

This project follows a Design Science Research (DSR) approach, which is commonly used for the iterative development and evaluation of sociotechnical systems. DSR emphasizes the creation of artifacts grounded in real-world problems and evaluated through utility, usability, and potential impact [12, 13]. Given the emotionally sensitive context of displacement narratives and the early-stage nature of this work, we adopted a formative expert evaluation strategy for initial feedback.

The prototype was informed by an earlier qualitative participatory study involving 12 individuals with forced migration experiences [14]. Participants reflected on what kinds of stories they would want to revisit or "replay" in a game, as well as on emotional needs, empowerment, and audience takeaways. These insights guided the interviewer's tone, prompt style, and scene design. The AI chatbot questions were adapted from themes and responses identified in that prior study, so that participant perspectives shaped the system from the outset.

At this stage, our primary goal was to assess whether the AI-powered interviewer could provide a low-pressure, emotionally neutral environment for self-narration. To do so, we conducted informal sessions with two expert participants, followed by semi-structured feedback on emotional comfort, conversational flow, and self-reflection. Their feedback was used to refine the system's prompting behavior and to assess the feasibility of replayable scenes as tools for narrative reflection prior to future community-informed evaluations.

4. System Overview and Implementation

Our system is designed as a modular, trauma-aware storytelling pipeline. In the current prototype, it consists of five components: introduction and onboarding, a locally hosted chatbot interview, narrative extraction into a structured scene representation, a prototype scene replay interface, and post-game

reflection. The core interview, extraction, and replay pipeline runs locally to preserve privacy and narrative control. An optional visual-context module can call an external image-generation API, but this feature is disabled by default and is not required for the interview-to-scene workflow reported in this paper. We use "structured scene representation" to refer to the extracted fields (schema) and dialogue prompts derived from the interview log. Figure 1 summarizes the current pipeline.

The interviewer is implemented as a Gradio-based local web interface connected to a locally hosted LLM served through LM Studio. During the interview, the system maintains the conversation history and applies a fixed interviewer prompt that instructs the model to ask one open-ended question at a time, avoid judgmental or therapeutic language, and remind participants that they may skip questions or stop at any time. After the session, the transcript is stored locally and passed to a post-processing step that extracts a compact scene representation in JSON format. This intermediate representation is then converted into scene data used to initialize the replay prototype.

The interviewer prompt was designed as a structured, trauma-aware questioning scaffold rather than a free-form conversational agent. Its role is to encourage self-paced narration while minimizing social pressure. The prompt instructs the model to use short, open-ended follow-up questions, avoid leading interpretations, and not present itself as a therapist or counselor. The wording of the prompts was informed by themes identified in our earlier participatory study, especially around belonging, identity, meaningful memories, and what participants might want to revisit or reframe through interactive scenes.

In the current prototype, narrative extraction is implemented as a schema-filling step over the completed interview transcript. The system produces five fields: `setting`, `character`, `memory`, `start_prompt`, and `prompt_template`. Here, `setting` captures the scene context, `character` describes the focal perspective, `memory` summarizes the core remembered situation, `start_prompt` initializes the replay scene, and `prompt_template` conditions subsequent turns. This intermediate JSON representation is intentionally lightweight: it makes the transition from interview text to interactive scene explicit, inspectable, and easy to revise in later iterations.

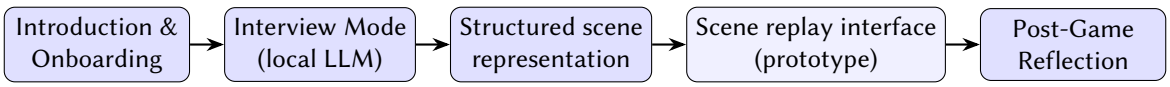

Figure 1: System overview of the current prototype. The evaluation in this paper focuses on the interview experience and the early narrative-to-scene flow.

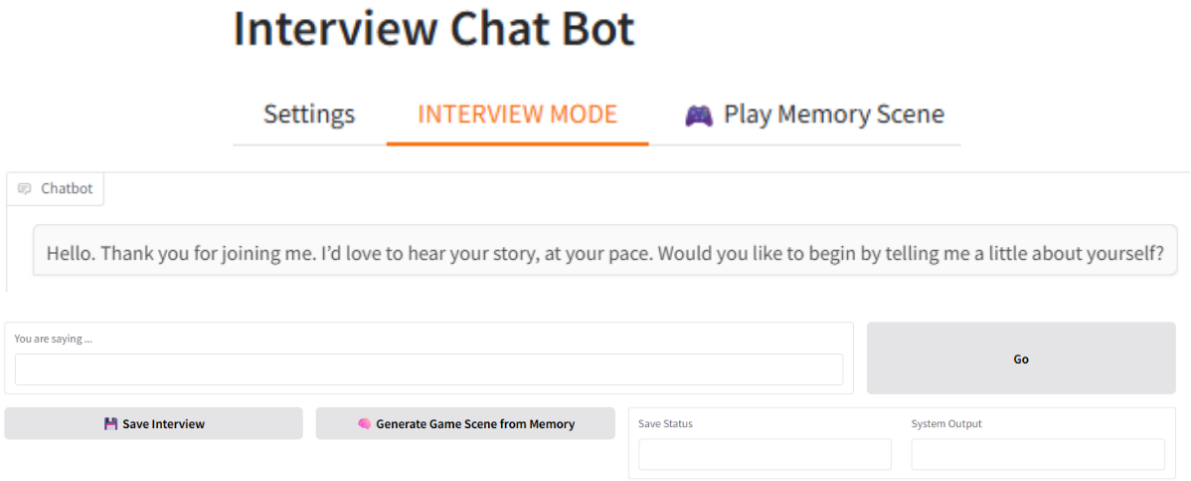

Figure 2: Interview interface of the current prototype. The system guides participants through self-paced narrative elicitation and allows them to save the interview and generate a replayable scene representation.

Introduction and Onboarding Participants begin with a short onboarding step that explains the project goals, the local processing approach, and the voluntary nature of participation. The onboarding emphasizes emotional safety: participants can skip questions, stop at any time, and avoid topics they do not wish to revisit. The system is framed as a research prototype for reflective storytelling, not a therapeutic tool.

Interview Mode A locally running LLM (*Zephyr* [15]) powers a Gradio-based chatbot [16] that guides participants through a reflective conversation. The interaction is designed to be emotionally neutral and low-pressure. Participants can answer briefly or in depth, and can skip prompts or end the session at any time. The interviewer structure is informed by themes and question styles derived from prior participatory input. Prompts focus on experiences of identity, belonging, and meaningful memories that participants may later want to revisit or reframe through an interactive scene. Figure 2 shows an example interaction in the interview interface.

Structured Scene Representation After the interview, the system parses the conversation log and extracts a structured scene representation. In the current version, the extracted representation includes five elements: `setting`, `character`, `memory`, `start_prompt`, and `prompt_template`. This representation serves as an intermediate format that supports replay and later extensions.

Scene Replay Interface (Prototype) The extracted scene representation is used to initialize a lightweight replay interface. In the current prototype, interaction is implemented as guided replay and continuation: the `start_prompt` initializes the scene, and the `prompt_template` conditions subsequent turns. Participants can continue the narration, test alternative actions, add clarification, or rephrase details as the scene progresses. This prototype is intended to demonstrate the transition from interview data to an interactive replay format rather than a fully developed branching game.

5. Formative Expert Evaluation

To assess perceived comfort, emotional safety, and usability of the AI interviewer, we conducted a small formative evaluation with two expert participants. Each participant engaged with the chatbot for approximately 10–15 minutes and then completed a semi-structured feedback interview. We use the term *expert participants* to refer to domain-informed evaluators selected for early-stage risk-screening rather than target-population validation.

While the system is ultimately intended for individuals with lived experiences of displacement, we began with expert participants without such lived experience as an initial safety and usability step. Neither participant reported lived experience of forced migration or displacement; however, both had personal experience of international relocation, which shaped how they related to themes of identity and belonging raised in the interview. One expert had a background in AI and game development, and the other had expertise in cultural studies and trauma-aware media, including professional experience working with marginalized populations. This early-stage evaluation was designed to identify usability issues and potential emotional risks before future community-informed studies with the intended participant groups. This approach is consistent with trauma-informed computing, which emphasizes evaluation practices that prioritize user safety and well-being in sensitive contexts [17].

Participants were asked to recount a meaningful memory via the chatbot. After the session, they responded to three guiding questions:

- Did you feel comfortable speaking to the AI?
- Would you say the same things to a human interviewer?
- Did you feel like the AI misunderstood you?

Responses were reviewed qualitatively alongside interaction logs to inform iterative refinement rather than to support generalizable claims.

Overall, both participants described the interviewer as reflective and low-pressure. One participant characterized the experience as "therapeutic" and noted that "nobody ever asks these questions," while the other appreciated the early conversational flow but found that it became more "bot-like" over time.

Feedback also highlighted concrete areas for improvement. Participants wanted better handling of long responses and a dialogue style that acknowledges or summarizes what was said (e.g., brief confirmation or recap). Minor usability issues were also noted, including input behavior and limited ability to review prior entries. Although neither participant had lived forced-migration experience, both connected the prompts to their own experiences as immigrants and reflected on themes of belonging. Given the exploratory nature of this expert study (n=2), we use these findings to guide iterative design improvements and refine the protocol for future community-informed evaluations.

6. Discussion and Future Work

This work presents an early-stage pipeline for trauma-aware storytelling centered on a locally hosted, emotionally neutral AI interviewer. The interviewer is intended to reduce social pressure and support participant control through self-paced disclosure, optional skipping of prompts, and the ability to stop at any time.

Findings from the formative expert sessions suggest that the perceived safety of an AI interviewer depends not only on topic selection and tone, but also on interaction behavior. Participants valued the reflective nature of the questions, yet highlighted that limited acknowledgement of long answers and occasional "forced turns" could undermine a sense of being heard. This points to the importance of conversational behaviors such as summarizing longer responses, confirming key details, and using targeted follow-up questions to preserve coherence and participant trust.

The current prototype also demonstrates a concrete transition from interview logs to a structured scene representation and a lightweight replay interface. While the replay component is not yet a full game, it shows how participant-authored memories can be transformed into an interactive format for revisiting and rephrasing narrated experiences. At the same time, this early version still requires stronger grounding and a clearer interaction model to better support participant agency during replay.

This paper reports an early-stage prototype and a small expert evaluation (n=2). The system has not yet been evaluated with participants who have lived experiences of displacement; this staging decision reflects an emphasis on risk screening, ethical approvals, and community-informed processes before engaging with potentially affected groups. The current findings should therefore be understood as formative and exploratory rather than generalizable.

Future work will (1) clarify and expand the interaction model beyond guided replay, (2) improve turn-taking behavior through acknowledgement and summarization of longer responses, and (3) address known LLM constraints such as handling emotionally nuanced inputs, occasional hallucinations, and limited conversational memory through guardrails and lightweight memory mechanisms. Automated memory elicitation also raises ethical risks, including distress during recall, over-disclosure, and misinterpretation. We therefore emphasize participant control, transparent framing as a research prototype rather than a therapeutic tool, and staged evaluation. The core interview, extraction, and replay pipeline runs locally, supporting privacy and data sovereignty by avoiding cloud-based processing in these stages.

7. Conclusion

This paper introduced a work-in-progress pipeline for trauma-aware, participatory storytelling using locally hosted AI. The system centers an automated interviewer and transforms interview logs into a structured scene representation that can be replayed through a prototype interface. We position replay as a design mechanism for revisiting and rephrasing narrated memories, while prioritizing privacy, consent, and participant control through local processing and staged evaluation.

Although motivated by stories of migration and displacement, the approach may be applicable to other sensitive contexts where users benefit from self-paced narration and careful handling of personal memories (e.g., grief, chronic illness, and identity-related experiences). Future work will refine conversational turn-taking (e.g., acknowledgement and summarization), clarify the interaction model beyond guided replay, and continue community-informed evaluation practices. We hope this direction supports more agency-respecting ways of documenting lived experience.

Declaration on Generative AI

During the preparation of this manuscript, the author(s) used a generative AI tool for language editing (grammar and clarity). The author(s) reviewed and revised the resulting text and take full responsibility for the content of the publication.

References

[1] C. Spurgeon, Participatory media and 'co-creative' storytelling, Media International Australia 154 (2015) 132–138. URL: https://doi.org/10.1177/1329878X1515400116. doi:10.1177/1329878X1515400116. arXiv:https://doi.org/10.1177/1329878X1515400116.

[2] A. Serbanescu, Human-ai system co-creativity to build interactive digital narratives, in: Lecture Notes in Computer Science, 2023. doi:10.1007/978-3-031-49811-4_37.

[3] A. Wuttke, M. Aßenmacher, C. Klamm, M. M. Lang, Q. Würschinger, F. Kreuter, AI conversational interviewing: Transforming surveys with LLMs as adaptive interviewers, in: A. Kazantseva, S. Szpakowicz, S. Degaetano-Ortlieb, Y. Bizzoni, J. Pagel (Eds.), Proceedings of the 9th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature (LaTeCH-CLfL 2025), Association for Computational Linguistics, Albuquerque, New Mexico, 2025, pp. 179–204. URL: https://aclanthology.org/2025.latechclfl-1.17/. doi:10.18653/v1/2025.latechclfl-1.17.

[4] E. Kim, J. Suk, S. Kim, N. Muennighoff, D. Kim, A. Oh, LLM-as-an-interviewer: Beyond static testing through dynamic LLM evaluation, in: W. Che, J. Nabende, E. Shutova, M. T. Pilehvar (Eds.), Findings of the Association for Computational Linguistics: ACL 2025, Association for Computational Linguistics, Vienna, Austria, 2025, pp. 26456–26493. URL: https://aclanthology.org/2025.findings-acl.1357/. doi:10.18653/v1/2025.findings-acl.1357.

[5] Y. Liu, H. Al-Sahaf, A. Chen, Y. Mei, M. Zhang, B. Xue, Pangea: Procedural artificial narrative using generative AI for turn-based games, arXiv preprint arXiv:2404.19721 (2024). URL: https://arxiv.org/abs/2404.19721.

[6] T. Rist, Using a large language model to turn explorations of virtual 3d-worlds into interactive narrative experiences, in: 2024 IEEE Conference on Games (CoG), IEEE, 2024, pp. 1–8.

[7] T. P. Hunt, Figs, A. France, Bury me, my love, 2017. Video game.

[8] C. Creations, Path out, 2017. Video game.

[9] L. Chouliaraki, Symbolic bordering: The self-representation of migrants and refugees in digital news, Popular Communication 15 (2017) 78–94. doi:10.1080/15405702.2017.1281415.

[10] M. Flanagan, Critical Play: Radical Game Design, MIT Press, Cambridge, MA, 2009.

[11] S. Danilovic, K. Chee, M. Skop, Playful resilience: Empowering recovery through autobiographical game-based storytelling in the opioid epidemic, in: Proceedings of the ACM on Human-Computer Interaction, CHI PLAY, volume 8, 2024, pp. Article 330, 35 pages. doi:10.1145/3677095.

[12] A. R. Hevner, S. T. March, J. Park, S. Ram, Design science in information systems research, MIS quarterly (2004) 75–105.

[13] K. Peffers, T. Tuunanen, M. A. Rothenberger, S. Chatterjee, A design science research methodology for information systems research, Journal of management information systems 24 (2007) 45–77.

[14] K. Harshina, et al., Experiencing forced migration: A participatory approach to developing a

serious video game, in: Games and Learning Alliance (GALA) Proceedings, Springer Nature Switzerland, 2024, pp. 388–393.
[15] L. Tunstall, E. Beeching, N. Lambert, N. Rajani, K. Rasul, Y. Belkada, S. Huang, L. Von Werra, C. Fourrier, N. Habib, et al., Zephyr: Direct distillation of lm alignment, arXiv preprint arXiv:2310.16944 (2023).
[16] Gradio, Gradio: Programmatic Interface for Building Machine Learning UIs, https://www.gradio.app/, 2025. [Online; accessed 30-May-2025].
[17] J. X. Chen, A. McDonald, Y. Zou, E. Tseng, K. A. Roundy, A. Tamersoy, F. Schaub, T. Ristenpart, N. Dell, Trauma-informed computing, in: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, CHI '22, Association for Computing Machinery, New York, NY, USA, 2022. URL: https://doi.org/10.1145/3491102.3517475. doi:10.1145/3491102.3517475.

A. Example Prompt and Scene Representation

Interviewer prompt excerpt. The interviewer prompt is designed as a gentle, trauma-aware scaffold. It instructs the model to ask one question at a time, avoid invasive or judgmental language, allow the participant to skip or pause, and maintain a calm, curious tone.

> Ask one question at a time.
> Be specific, but not invasive.
> Allow the participant to skip or pause at any moment.
> Always affirm, never judge.
> Keep your tone gentle and curious.
> Never offer advice or opinions.

Extraction prompt excerpt. After the interview, the transcript is passed to a local extraction step that prompts the model to transform the conversation into a structured replayable scene.

> Your task is to turn an interview into an interactive narrative scene.
> Focus on a story the participant shared where they wished things had gone differently.
> Reframe the memory as an interactive scene the participant could replay with more emotional agency.
> Return a JSON object with these fields: setting, character, memory, prompt_template, start_prompt.

Example structured scene representation. The following synthetic example illustrates the intermediate representation used to initialize replay:

```
{
  "setting": "A hometown revisited through memory and longing",
  "character": "You return carrying grief, nostalgia, and uncertainty",
  "memory": "You wish to revisit a place that no longer feels fully reachable",
  "start_prompt": "You stand on the old street again. What do you do first?",
  "prompt_template": "..."
}
```