Explore the Future Earth with Wander 2.0:
AI Chatbot Driven By Knowledge-base Story Generation and Text-to-image Model

Yuqian Sun*
yuqiansun@network.rca.ac.uk
Computer Science Research Centre,
Royal College of Art
London, UK

Ying Xu
yingxu2030@gmail.com
Department of Industrial Design,
School of Art and Design, Wuhan
University of Technology
Wuhan, China

Chenhang Cheng
koch3092@gmail.com
Unaffiliated
Shanghai, China

Yihua Li
cory958014884@gmail.com
Department of Product Design,
College of Fashion and Design,
Donghua University
Shanghai, China

Chang Hee Lee
changhee.lee@kaist.ac.kr
Affective Systems and Cognition Lab,
Industrial Design Department,
College of Engineering, Korea
Advanced Institute of Science and
Technology
Daejeon, South Korea

Ali Asadipour
ali.asadipour@rca.ac.uk
Computer Science Research Centre,
Royal College of Art
London, UK
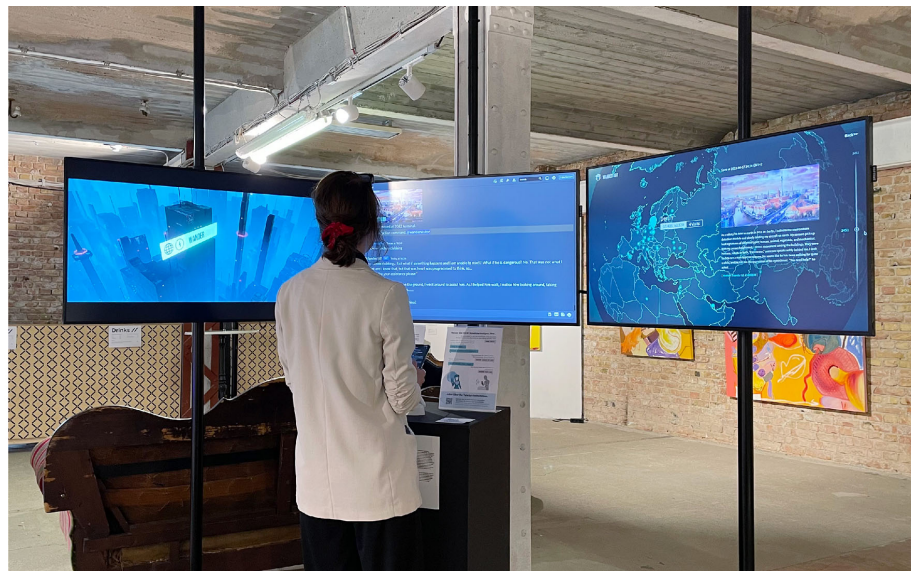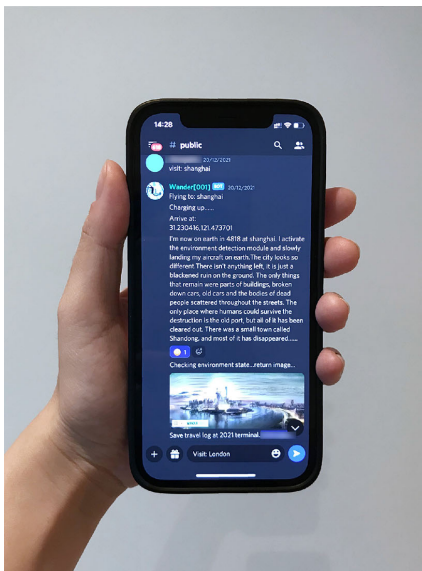
Figure 1: (Left) Wander bot on phone (Right) Exhibition photo.

*Corresponding author

ABSTRACT

People always envision the future of earth through science fiction (Sci-fi), so can we create a unique experience of "visiting the future earth" through the lens of artificial intelligence (AI)? We introduce Wander 2.0, an AI chatbot that co-creates sci-fi stories through knowledge-based story generation on daily communication platforms like WeChat and Discord. Using location information from Google Maps, Wander generates narrative travelogues about specific locations (e.g. Paris) through a large-scale language model (LLM). Additionally, using the large-scale text-to-image model (LTGM) Stable Diffusion, Wander transfers future scenes that match

both the text description and location photo, facilitating future imagination. The project also includes a real-time visualization of the human-AI collaborations on a future map. Through journeys with visitors from all over the world, Wander demonstrates how AI can serve as a subjective interface linking fiction and reality. Our research shows that multi-modal AI systems have the potential to extend the artistic experience and creative world-building through adaptive and unique content generation for different people. Wander 2.0 is available at http://wander001.com/

CCS CONCEPTS

• **Applied computing** → **Media arts**; • **Human-centered computing** → *Interaction design process and methods*; **Text input**.

KEYWORDS

Artificial intelligence, human-AI interaction, chatbot, gaming, interactive fiction, design fiction

1 INTRODUCTION

How will the future earth look like through an AI agent's eyes? This paper presents a narrative AI chatbot called Wander, which uses knowledge-based story generation to produce interactive fiction on daily communication platforms such as Discord[1]. Relying on the users' language text inputs, Wander creates science fiction stories based on real-world map information and photographs, as well as the future scene generated by text-to-image Stable Diffusion model.

This work aims to capture the sense of "visiting the future earth" through an approach of human-AI collaboration. Each time having received a location message sent by the participant, Wander generates a unique, science fiction-style interactive travelogue based on real-world location information and photos from the Google Knowledge Graph[6]. Through free input text commands without predefined options, participants can direct Wander to explore locations on the "future earth". The records of these travelogues are visualized on a future global map, which updates in real-time through public participation. The interpretation of outcomes varied through the audience's personal choices, memories, and expectations.

In the previous version of Wander 1.0 [15], the authors discussed how people would interact with an AI agent in a narrative context of the future metaverse. In this upgraded version, with the state-of-the-art large-scale text-to-image generation model (LTGM) Stable Diffusion[5], Wander can transfer the destination photo into a futuristic style while keeping the generated image's content consistent with the text. By taking this approach, this work provided collaborative speculation[11] to how human may co-explore a concept of " future earth" with AI.

Inspired by text-based adventure game like Multi-player dungeon (MUD) [16] and prior study that enabled a collaborative experience in which users' fantasies were facilitated by AI-generated visuals [11], this interactivity provides a crosspoint between reality and virtuality and turns data of familiar surroundings into future speculations. Moreover, with the integration of a story generation model, state-of-the-art LTGM, together with knowledge graph with a visualization map, we contributed a practical workflow on how to develop an AI-based collaborative chatbot that can serve as an intelligent storyteller.

2 CONCEPTUAL BACKGROUND

The idea of ''future traveller" took shape during the COVID-19 pandemic when traveling was limited. Our work was initially motivated by an observation between the travel limitations and numerous advancements of generative AIs which have shown great potential to facilitate creativity. We aim to articulate if there is any opportunity for human to leverage the capability of generative AI and has an "imaginary travel" with AI that is beyond time and space. Just like UC Berkeley's graduation ceremony in Minecraft [2] and Travis Scott's concert in Fortnite [3], the virtual world extends people's experience of daily life beyond the limits of geographical location and the pandemic. Everyone can easily participate in it. Generally, people have always used science fiction as a means of imagining the future. While leveraging the features of AIGC nowadays, people can imagine the future, which looks different but somewhat familiar.

3 RELATED WORKS

We focus primarily on projects using large language models (LLMs) which can achieve dynamic content generation beyond pre-written rules and contents. Through this, AI agents are becoming intimate companions in Replika [7], social members in Xiaoice Island [4], and even actors in talk shows [3]. More recently updated dialogue models, such as ChatGPT [12] and Character.ai [10], have even more capabilities, ranging from customizing systems (simulating a library or operating system, writing fiction, and even papers) to playing roles.

However, besides the progress in LLM, researchers questioned the agency and autonomy of models since the generated content is backward-oriented through the collection of data [13]. We hope to mediate this argument through an artistic experience. Even though AI systems generate outcomes through the analysis towards data, the meaning for humans is left for subjective interpretations. Researchers in HCI provides a framework and metrics for "meaningful interaction" [9], but still cannot evaluate it since it's subjective. This leaves space for artists and designers to consider how human-AI interactions can become meaningful. Our entry is to introduce real-world information, like map, which can connect and resonate to public viewers.

We got inspiration from text adventure games like Multi-player Dungeon (MUD) [16], wherein players use text descriptions to choose the next step, and prior work in which AI-generated visuals were used to enhance users' imagination on utopian [11]. This type

---

of experience has also been extended through LLM in recent works like AI Dungeon [1] and 1001 Nights[14], in which players can use free inputs without fixed choices. In this way, human-AI interaction can make sense of uniqueness and meaning for each person, drawing upon their varied memories and experiences, instead of the formalistic plots the developer presented in advance. In this work, we developed an AI-based chatbot that served as a narrator between the fictional and the real world, allowing people to visit their hometowns or dream travel destinations through the lens of future adventure.

4 IMPLEMENTATION

Our approach contains two parts: Wander chatbot (conversational agent) and the future map. During the exhibition(Fig. 4) , we put three screens for the installation: (1) Concept Video (2) Chatbot Channel (3) Future Map. The audience can use the iPad to send messages or scan the QR code to join by phone. When Wander sends back the initial travelogue, the map will automatically rotate and pop-up the record. We imitate the experience as how people search locations on Google Maps, to bring the feel of "see the future earth".

4.1 Wander Chatbot

The core part of this project is a conversational agent called "Wander". According to the background story we set, She is an android that explores and collects data from the future earth. She needs human's help to visit different places, and She can receive commands from the 21st century through Discord.

Each time when a participant sends a location message with "!Visit" command, She will get to that place in a random year between 2600 and 5000 AD, then sends back travel notes including GPS location, environment description and an image. The text description is real-time generated based on the knowledge graph of the location. Then through text-to-image model Stable Diffusion, the actual photo from Google Map will turn to a future version that matched the description. Then, with the "!action" command, participants can ask Wander to explore the place in any method, like searching for lives, getting into the ruins, etc.

4.2 Future Map

Future map is a visualisation website for public participation. After receiving the generated results of "!visit" command, the future map will rotate and scale to the destination, then this travelogue will update and show up (Fig. 4). Travelogues in each location would be sorted in the timeline. Each location has a lighting pillar that visualised the number of visits. Since "!action" commands may contain private preference, only records of "!Visit" commands would be shown.

This experience is also accessible out of the exhibition. Participants can travel with Wander on their phone or PC through discord channel and the website. People can rotate the globe and check the journey histories of each location.

4.2.1 Methods.

Story generation model. The text generation model we used is called dreamily.ai [8], a creative writing platform using a modified transformer (a self-attention multi-layer neural network) model

Kashmir

Year 4142 / **GPS** 33.277839,75.34121789999999

Knowledge Graph

Kashmir is the northernmost geographical region of the Indian subcontinent [...]

Story Generation

It's not hard to imagine the story behind this phenomenon. After all, the earth has been affected by alien technology. It had become more advanced than the human world, and humans have become extinct. Now, it is just a place filled with **dust and debris**. After a moment, I get down from my plane. I check the weather conditions. There is no visible sign of **solar energy** in the sky, although there might be an **atmospheric layer of clouds** above us......

Birmingham

Year 4710 / **GPS** 52.48624299999999,-1.890401

Knowledge Graph

Birmingham is a city and metropolitan borough in the metropolitan county of West Midlands in England. It is the second-largest city [...]

Story Generation

A large **orange dome** was placed over the whole country of west Midlands, protecting it from the radiation from the meteorites. While landing on the ground I take out one of my microphones and scan for any signals indicating danger. When I confirm that there's no danger detected, I begin moving around the area observing its surroundings and mapping the environment of the place. The sky has become purple after several hours......
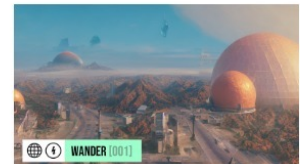
Figure 2: Example travelogues from audience interactions.

Yuqian Sun, Ying Xu, Chenhang Cheng, Yihua Li, Chang Hee Lee, and Ali Asadipour
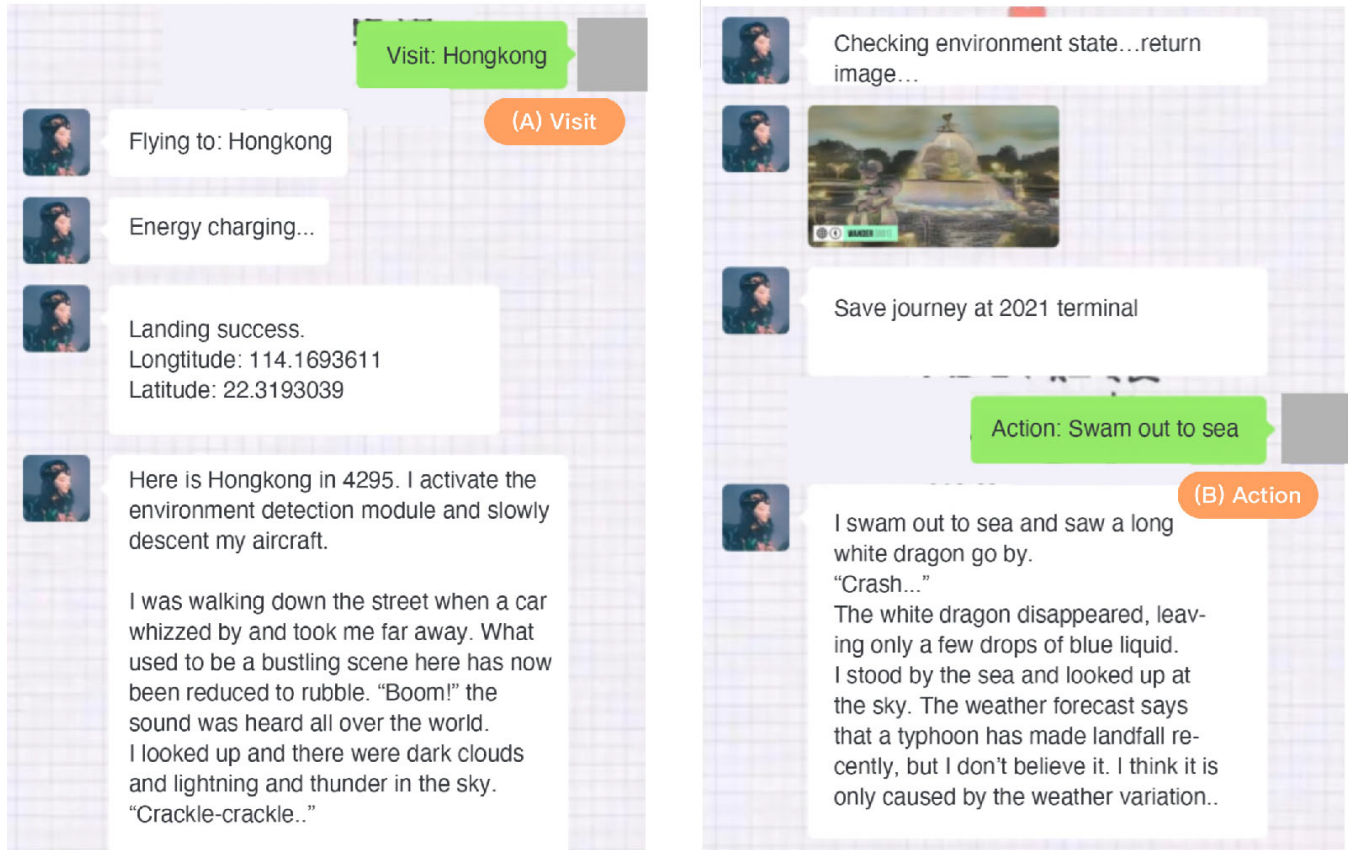

Figure 3: Conversation with Wander.

trained with quality fiction. Both the English and Chinese data sets consist of open access to fan-fiction and e-books on the web, and both data set sizes are about 100 GB. Although this model is not appropriate for ubiquitous tasks (e.g. writing official documents or codes), it is suitable for story generation. We used its zero-shot story generation ability with a designed prompt. The prompt of Wander includes two parts: (1) Wander's self-introduction, including her motivation and task in the future world and (2) Summarised description of the visited location. We used Google Knowledge Graph to extract descriptive sentences into a short introduction of the location. With the location-specific prompt, Wander will produce a rich and surprising response that corresponds to the destination. When Wander sends back the initial travelogue, participants can send "!action" command (e.g. "search for life") to instruct Wander to explore the location. The message will be interpreted as first-person action of Wander ("I search for life,"), added to the prompt, and generated the result. The response will be saved in the prompt for the next "!action" command, and so on. Due to the memory limitations of the model, Wander is limited to only 8 actions.

Image generation. In this version, for image generation, we used stable diffusion [5], a latent diffusion model conditioned on the (non-pooled) text embeddings of a CLIP ViT-L/14 text encoder [2]. This model is able to command image-to-image tasks through

text prompts, which means, we can modify real-world photos on Wander's description.

Through Google Maps' API, we can get the GPS location and the image of the place, if it exists. We designed a realistic concept-art style prompt for image generation, and the corresponding parts will be replaced by the name of the destination and Wander's description. After experiments, we put the strength parameter to 0.7 with 30 steps to balance the original photo and the diffusion generation. Through this real-time workflow, the future scene will generally transfer Wander's observations of the place into a future image that also keeps the features of the original photo. The total generation task will take around twenty seconds on A100 GPU.

5 CONCLUSION

In this design fiction work, chatbot Wander utilizes knowledge-based story generation and a text-to-image model to create interactive fiction on daily communication platforms. As a creative system, She demonstrates procedural playable story generation in a speculative situation. As a storyteller, She became a subjective interface that mediates human-AI relationships through explorations on a future earth that is generated through map information. Up until 2023 January, Wander met over 20000 visitors, who left feedback on how an accessible AI narrator, similar to characters in video games, can tell meaningful stories together with humans. We believe that in the near future, artistic experience and world-building can be

Figure 4: (Top) Exhibition and (Bottom) future map.

designed by individuals, conveyed and extended by multi-modal AI systems, and interpreted by the audience.

ACKNOWLEDGMENTS

REFERENCES

[1] 2020. AI Dungeon. https://play.aidungeon.io/main/home.
[2] OpenAI 2020. *CLIP*. OpenAI. https://github.com/openai/CLIP
[3] 2020. Robot Bores: AI-powered awkward first date. *BBC News* (Nov. 2020). https://www.bbc.com/news/technology-54718671
[4] 2021. Microsoft Xiao Ice Land. https://island.xiaoice.com/
[5] 2022. *CompVis/Stable-Diffusion · Hugging Face*. https://huggingface.co/CompVis/stable-diffusion
[6] 2022. Google Knowledge Graph Search API. https://developers.google.com/knowledge-graph
[7] 2022. Replika. https://replika.ai.
[8] Caiyun. 2021. Dreamily-beta. https://dreamily.ai/#/
[9] Vanessa Julia Carpenter and Elisa D. Mekler. 2019. Towards metrics of meaningfulness for Tech Practitioners. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (2019). https://doi.org/10.1145/3290607.3299060
[10] Character.ai. 2022. *Character.AI*. Character.AI. https://beta.character.ai/
[11] Ziv Epstein, Hope Schroeder, and Dava Newman. 2022. When happy accidents spark creativity: Bringing collaborative speculation to life with generative AI. *arXiv preprint arXiv:2206.00533* (2022).
[12] OpenAI. 2022. *ChatGPT: Optimizing Language Models for Dialogue*. OpenAI. https://openai.com/blog/chatgpt/
[13] Silvia Pierosara. 2022. Narrative autonomy and artificial storytelling. *AI & SOCIETY* (2022). https://doi.org/10.1007/s00146-022-01595-9
[14] Yuqian Sun, Xuran Ni, Haozhen Feng, Ray LC, Chang Hee Lee, and Ali Asadipour. 2022. Bringing Stories to Life in 1001 Nights: A Co-creative Text Adventure Game Using a Story Generation Model. In *Interactive Storytelling: 15th International Conference on Interactive Digital Storytelling, ICIDS 2022, Santa Cruz, CA, USA, December 4–7, 2022, Proceedings* (Berlin, Heidelberg, 2022-12-04). Springer-Verlag, 651–672. https://doi.org/10.1007/978-3-031-22298-6_42
[15] Yuqian Sun, Ying Xu, Chenhang Cheng, Yihua Li, Chang Hee Lee, and Ali Asadipour. 2022. Travel with Wander in the Metaverse: An AI chatbot to Visit the Future Earth. In *2022 IEEE 24th International Workshop on Multimedia Signal Processing (MMSP)*. 1–6. https://doi.org/10.1109/MMSP55362.2022.9950031
[16] R Trubshaw and R A Bartle. 1978. *Multi-user dungeon*.