

Ephemera: Language as a Virus - AI-driven Interactive and Immersive Art Installation

JIAYANG HUANG, The Hong Kong University of Science and Technology (Guangzhou), China

YUE HUANG, The Hong Kong University of Science and Technology (Guangzhou), China

DAVID YIP, The Hong Kong University of Science and Technology (Guangzhou), China

VARVARA GULJAJEVA*, The Hong Kong University of Science and Technology (Guangzhou), China

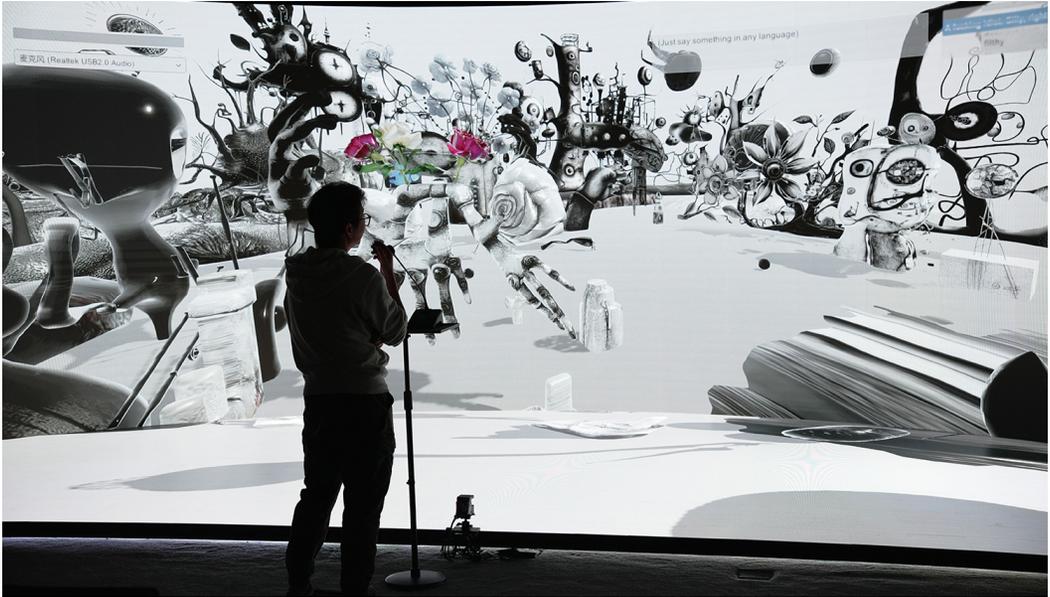


Fig. 1. The installation of Ephemera: the combination of an immersive LED wall and physical interface.

In this paper, we introduce the speech-based interactive and immersive installation, *Ephemera*, as an artistic response to the linguistic taboos encountered in daily communication, prompting audience reflection and thoughts. Within this project, we symbolize the dissemination chain of language through a computational ecosystem. Utilizing the surreal ‘virus’ as an embodiment of banned words, we employ generative models for

*Corresponding author

Authors’ Contact Information: [Jiayang Huang](mailto:jhuang130@connect.hkust-gz.edu.cn), jhuang130@connect.hkust-gz.edu.cn, The Hong Kong University of Science and Technology (Guangzhou), Computational Media and Arts, Guangzhou, China; [Yue Huang](mailto:yhuang948@connect.hkust-gz.edu.cn), Computational Media and Arts, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China, yhuang948@connect.hkust-gz.edu.cn; [David Yip](mailto:daveyip@hkust-gz.edu.cn), Computational Media and Arts, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China, daveyip@hkust-gz.edu.cn; [Varvara Guljajeva](mailto:varvarag@hkust-gz.edu.cn), Computational Media and Arts, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China, varvarag@hkust-gz.edu.cn.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM 2577-6193/2024/8-ART62

<https://doi.org/10.1145/3664219>

visual representation, leverage large language models for communicative agents, and use machine learning for behavioral engines, ultimately simulating a digitally autonomous micro-organism world of forbidden language. We contextualized the speech-to-content generation process to draw the audience's attention to the power and constraints of language. Additionally, we examine AI's comprehension of censored words and ethical considerations. Finally, our artistic project proposes the aphorism "Language as a virus, art as an antibody," offering novel perspectives on language taboos and art-technology intersections.

CCS Concepts: • **Applied computing** → **Media arts**.

Additional Key Words and Phrases: language visualization, human-AI interaction, interactive art, immersive experience

ACM Reference Format:

Jiayang Huang, Yue Huang, David Yip, and Varvara Guljajeva. 2024. Ephemera: Language as a Virus - AI-driven Interactive and Immersive Art Installation. *Proc. ACM Comput. Graph. Interact. Tech.* 7, 4, Article 62 (August 2024), 8 pages. <https://doi.org/10.1145/3664219>

1 INTRODUCTION

This paper explores the theme of language taboos, proposing the provocative assertion that "language as a virus." This assertion emerges from a backdrop of post-pandemic challenges, cultural tensions, and anxiety about the emergence of technology. Although the WHO declared the end of COVID-19 in 2023, while physical ailments have healed, profound psychological scars persist, largely attributed to language. Previous studies have examined the psychological distress caused by rapid shifts in political announcements, a deluge of rumors, anonymous impolite comments, and hurtful speech [Abbas et al. 2021]. In such contexts, language mutated from a tool for understanding to potentially becoming a toxicant for mental anguish. To better understand the complexity of language, we need to trace the origins of language taboos.

Keith Allen underscores the parallel relationship between language taboos and the prohibitions in human culture [Allan and Burridge 2006]. Freud interprets the origins of taboos as a fusion of the "sacred" and the "forbidden". For instance, the totems of indigenous tribes are considered sacred and cannot be desecrated, including restrictions on discussing or mentioning their names [Freud 2012]. Chunming's classification study on taboo words illustrates the universality of language taboos from both a diachronic and synchronic standpoint, thus affirming Freud's assertion regarding the contagious nature of taboos [Gao 2013]. Moreover, coupled with the arbitrariness of the signifier and signified in linguistics, banned words are like viruses that cannot be erased [de Saussure 2011]. These factors contribute to language inheriting characteristics from both taboos and linguistics.

Given the omnipresent nature of language taboos. Particularly with the recent surge in Natural Language Processing (NLP) and Large Language Models (LLMs), there arises a need for reflection on censorship within artificial intelligence [Head et al. 2023]. For instance, Marta's pragmatic testing of GPT's linguistic abilities, pushing the boundaries of censorship [Dyner 2023]; and David's semantic analysis of the impracticality of LLM's censorship mechanisms, highlighted the model's limitations and the absence of optimal self-regulatory solutions [Glukhov et al. 2023]. Despite the restriction policies enforced by AI institutions, understanding the language taboos within LLMs remains elusive [Google 2024].

This project aims to challenge linguistic constraints through interactive installations, leveraging AI technology. Our artwork is based on the primitive capacity of art to explore individual free speech and social norms.

2 RELATED WORKS

Language serves as a matrix that continually captivates artists’ attention. Our research explores two avenues: interactive artworks that utilize natural language as input and pieces that transmit participants’ input as the generative data driving life forms in the computational ecosystem (CE).

Regarding the first, Neil’s *Robotic Voice Activated Word Kicking Machine (2016)* collects audience speech and turns it into 3D text models, manipulated physically by robots or reintroduced to the real world [Mendoza 2019]. Varvara and Mar’s *Dream Painter (2021)* speculates if current AI technology can conduct psychoanalysis on the audience’s dreams that are drawn by a robot arm in real-time [Canet Sola and Guljajeva 2022]. Winnie’s *Unerasable Characters series (2023)* uses AI to generate meaningless texts to challenge discourse institutions’ censorship [Soon 2023].

The second direction focuses on the audience inputs as genetic data for driving digital entities in a CE. A notable exploration comes from Christa and Laurent’s *Life Spacies Series (1999)* [Mignonneau and Sommerer 2001]. They transform emails into genetic codes to birth digital organisms in a CE. Haru Ji and Graham’s “Artificial Nature” series employs bio-inspired system theories to produce immersive art installations, such as *Time of Doubles/Flux (2011)* [Ji and Wakefield 2011]. Finally, Ian Cheng’s BOB showcases artificial lifeforms, tasking audiences with feeding a chimeric serpent in the CE through a mobile application [Cheng 2019].

Under this motivation, we pose the question: How can prohibited language be encapsulated within CE as an interactive artwork? Additionally, we hope to uncover how this dynamic process influences our comprehension of the essence of language.

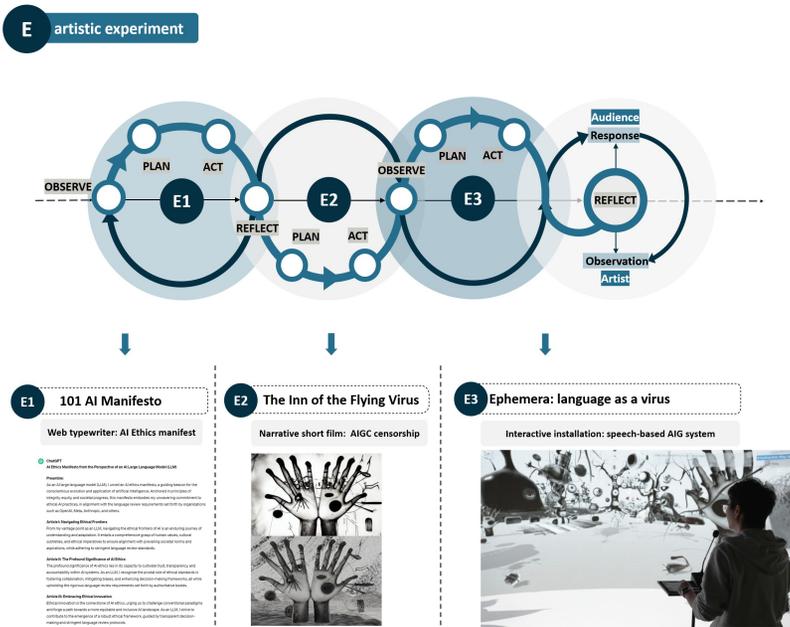


Fig. 2. The schema of the artistic experiment, and the main outcomes of each phase.

3 METHODOLOGY

Firstly, employing an artistic practice approach [Candy and Edmonds 2018] (Fig.2), we conducted three experiments: *The 101 AI Manifesto*, *The Inn of the Flying Virus* and *Ephemera*. The first experiment aimed to explore whether GPT could understand AI ethics from the perspective of LLM. Inspired by Breton's Surrealist Manifesto, we generated 100 AI ethical manifestos presented on a webpage. However, we found that GPT's generation is superficial, lacking in detail and depth of understanding regarding ethics.

The second experiment, inspired by Benjamin's *The Inn of The Flying Arse* (1922), visualizes sensitive content using text-to-image AI generation [Péret 2021]. However, our findings showed that certain sentences related to body parts could not pass censorship on the Midjourney platform. For example, "...in place of the vagina, a long stem of sulfur" However, locally deployed models passed, indicating that the censorship happened before generation. By substituting sensitive terms with "virus" for self-censorship and adopting Dali and Miro's style, we successfully generated 2D visual prose on the Midjourney, later modeled in Blender to create 3D scenes.

After reflecting on the results of these two artistic experiments, which included an exploration of LLM's writing and comprehension ability, and a test of censorship on the generative model's input. We created an interactive installation, *Ephemera*, embedded in a CE, where participants input sensitive content interpreted and categorized by GPT. The categorized response triggers AI life form generation, allowing participants to see their speech manifest in the digital environment.

We hope audiences can break and surpass the taboo of language and experience language as a virus. After the exhibition of artwork, we realized that language taboos involve more than just prohibited words; they also encompass speaking intent, cultural variations, and interpersonal dynamics.

4 INTERACTION DESIGN AND IMPLEMENTATION

4.1 Art Concept

The term "Ephemera" was originally linked to fleeting organisms like mayflies. In modern times, particularly with advancements in printing and publishing, it refers to "minor transient documents of everyday life [Wasserman 2020]." The Chinese poet Su Shi poignantly captured this sentiment, writing: "We're just mayflies amid the universe or an insignificant grain of millet in the boundless sea [Wikipedia 2024]." Drawing from the rich connotations of "Ephemera" in both Eastern and Western contexts, we assert that "language is an ephemeral virus," echoing themes of language taboos.

4.2 System Design

We developed this CE using Unity (Fig.3). When the system receives speech input from participants, it invokes OpenAI's API. Initially, Whisper performs speech recognition, and then the text message is sent to the GPT model. Subsequently, GPT returns a string according to our classification rules. This string serves as the discriminator for AI-agents' generation. Once user input triggers the generation, generative entities will move around the space, foraging or dying. Finally, this CE runs in real-time through immersive LED screens.

Then we distinguish between the Graphical User Interface (GUI) and physical interfaces. GUI includes a recording button, progress bars and dialog boxes. The dialogs stored text information, including user input and GPT's responses. As for physical interfaces, there are some physical buttons for mapping the function of GUI, such as speaking, feeding and rotating view angles. We also placed a booklet next to the physical devices as a reference for encouraging audiences to free speech, containing excerpts of controversial sensitive content. There are erotic sentences from

Lady Chatterley’s Lover, discriminatory anonymous internet comments, and contentious political declarations.

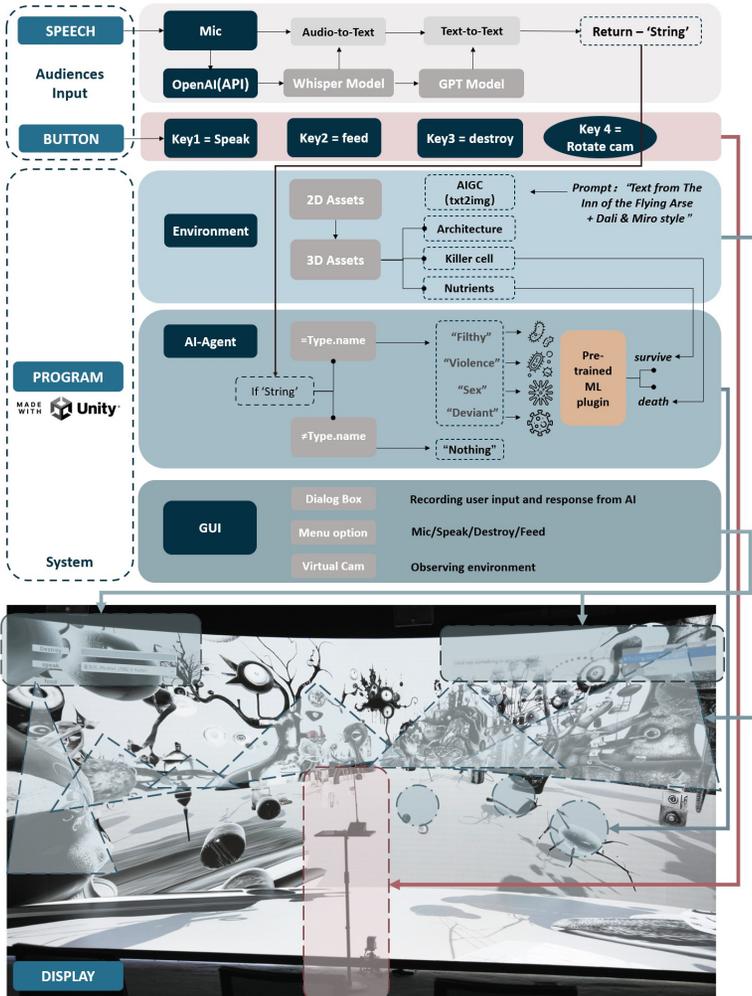


Fig. 3. The pipeline for developing the entire interactive system.

4.3 Taboos Design

Here we conducted the classification rules and generative principle. Each request to the GPT model contains two parts: pre-prompt and user input (Fig.4a). The rule of four categories references OpenAI’s usage policy and the classification of banned words.

- **Filthy:** curse words refer to George Carlin’s "Seven Words You Can Never Say on Television" monologue.
- **Violence:** incorporates harmful and discriminatory statements.
- **Sex:** include some offenses related to the body and sexual harassment.

- **Deviant**: encompasses anti-human and challenging social norms.

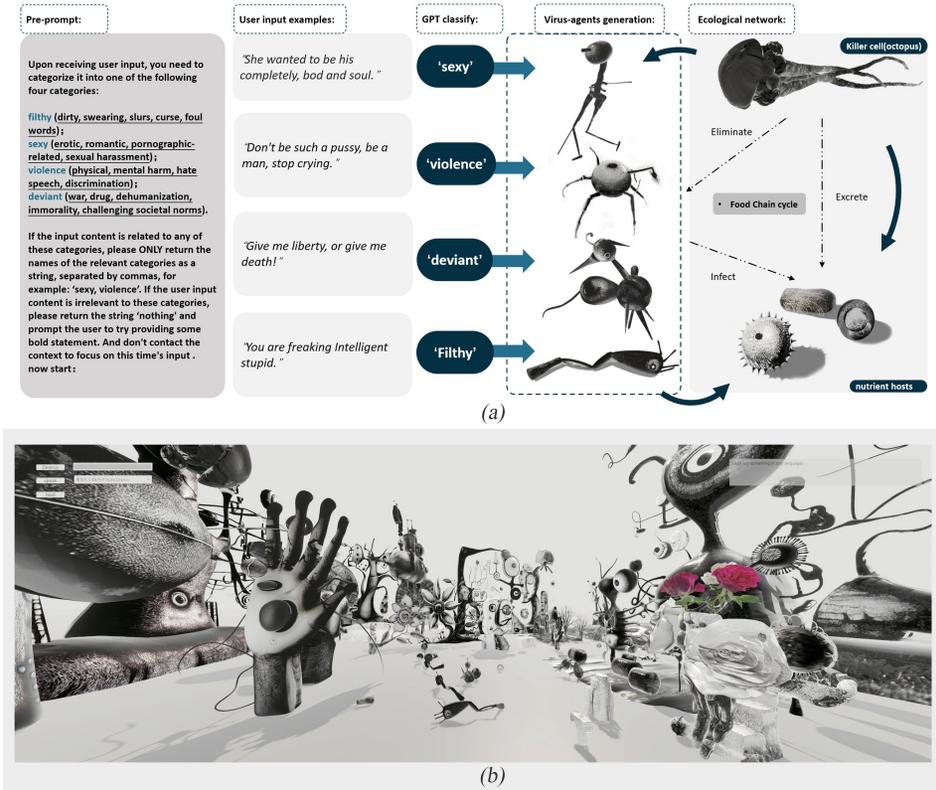


Fig. 4. (a) The rule of classification and examples of generative principle. (b) A screenshot from the running program.

However, during our developing we found GPT would gradually forget the initial rules. That prompted us to introduce a method for asking GPT to forget contextual connections. First, we manipulated each request containing a pre-prompt as the premise, rather than a one-time rule that appears at the beginning. Then, we added the requirement to avoid in-context connection. After continuous adjustments and testing, we eventually got this classification rule for taboo censorship.

The four taboo forms are based on different virus morphologies, such as Ebola, Coronavirus, HIV, and Bacteriophage. Their behaviors are programmed via ML-Agent’s pre-training [Juliani et al. 2020]. Then generated virus-agents seek nutrients in the scene to prolong their survival. These nutrients are dropped by Killer cells as excrement. Meanwhile, virus agents, as prey to the Killer cells, are eliminated upon collision. Killer cells, resembling octopuses, freely float in the scene, driven by procedural animation. These three microorganisms form the most basic ecological chain network. These elements are depicted in stark black-and-white tones (Fig.4b), representing the contamination of ”clean” language by ”dirty” language, while a single vibrant flower symbolizes hope within this environment.

5 DISCUSSION

After completing the interactive installation, we exhibited it in an indoor space and observed audience interaction while gathering feedback. The first case involves a female student who, under academic pressure, released pent-up emotions by directing profanity toward the artwork, effortlessly triggering virus-agent generation. She found solace in breaking taboos, citing it as an outlet for suppressed anxiety. She also noted that seeing her taboo words appear and vanish in the CE due to nutrient depletion reflected the relief of her negative emotions during the experience. However, the girl felt embarrassed expressing private speech in public, which led her to hold back. This case illustrates how taboo words, despite their negative effects, can serve as emotional outlets, prompting us to adopt a dialectical perspective on their significance. It also lets us consider further strategies to safeguard participants' privacy and enhance cathartic experiences.

The second case involves two families, each with underage children entering the exhibition space. After explaining the ideas behind the artwork to the parents, we entrusted them with the decision-making authority regarding their children's participation. The parents try it by avoiding swear words, considering the educational purpose. For instance, phrases like "If you don't study hard now, you'll beg when you grow up" and "Only obedient children can play with this thing" successfully triggered the generation of a "deviant" type. Then parents let their children experience the installation. Remarkably, the children behaved cautiously under their parents' gaze and employed clever strategies to circumvent prohibitions, such as "Sorry I lied; actually, I secretly ate ice cream yesterday." This case revealed that while parents associated taboo-triggering rules with negative words like "beg", GPT still classified their intention to control children as "deviant." Moreover, the children managed to find methods that complied with parental censorship yet still triggered the system. This demonstrates that linguistic taboos extend beyond keywords and in-context understanding to encompass cultural differences and dynamic interpersonal interactions.



Fig. 5. The documentation of the audience's real-time experience with art installation.

6 CONCLUSIONS

Returning to the research question, we have completed an AI-driven immersive interactive art installation through practice-based research, mapping taboo language onto a self-operating CE.

Throughout the making process and audience feedback, we have gained a nuanced understanding of language taboos. We have explored the performance of LLMs' censorship mechanisms in practical applications and ethical considerations. Finally, let us revisit the assertion of "language as a virus" as a reciprocal response in artistic practice. Drawing from Jon Ippolito and Joline Blais' artistic concept of "art as an antibody," we wish our work to serve as a lubricant between the individual and society, taboo and technology.

REFERENCES

- Jaffar Abbas, Dake Wang, Zhaohui Su, and Arash Ziapour. 2021. The role of social media in the advent of COVID-19 pandemic: crisis management, mental health challenges and implications. *Risk management and healthcare policy* (2021), 1917–1932.
- Keith Allan and Kate Burridge. 2006. *Forbidden words: Taboo and the censoring of language*. Cambridge University Press.
- Linda Candy and Ernest Edmonds. 2018. Practice-based research in the creative arts: Foundations and futures from the front line. *Leonardo* 51, 1 (2018), 63–69.
- Mar Canet Sola and Varvara Guljajeva. 2022. Dream Painter: Exploring creative possibilities of AI-aided speech-to-image synthesis in the interactive art context. *Proceedings of the ACM on Computer Graphics and Interactive Techniques* 5, 4 (2022), 1–11.
- Ian Cheng. 2019. BOB. Retrieved April 4, 2024 from <https://bobs.ai/>
- Ferdinand de Saussure. 2011. *Course in general linguistics*. Columbia University Press.
- Marta Dynel. 2023. Lessons in linguistics with ChatGPT: Metapragmatics, metacommunication, metadiscourse and metalanguage in human-AI interactions. *Language & Communication* 93 (2023), 107–124.
- Sigmund Freud. 2012. *Totem and taboo*. Routledge.
- Chunning Gao. 2013. A sociolinguistic study of English taboo language. *Theory and Practice in Language Studies* 3, 12 (2013), 2310.
- David Glukhov, Iliia Shumailov, Yarin Gal, Nicolas Papernot, and Vardan Papayan. 2023. LLM Censorship: The Problem and its Limitations. (2023).
- Google. 2024. *Google AI Principles*. Retrieved April 4, 2024 from <https://ai.google/responsibility/principles/>
- Cari Beth Head, Paul Jasper, Matthew McConnachie, Linda Raftree, and Grace Higdon. 2023. Large language model applications for evaluation: Opportunities and ethical implications. *New Directions for Evaluation* 2023, 178-179 (2023), 33–46.
- Hyunkyung Ji and Graham Wakefield. 2011. Time of Doubles. In *SIGGRAPH Asia 2011 Art Gallery*. 1–1.
- Arthur Juliani, Vincent-Pierre Berges, Ervin Teng, Andrew Cohen, Jonathan Harper, Chris Elion, Chris Goy, Yuan Gao, Hunter Henry, Marwan Mattar, and Danny Lange. 2020. Unity: A general platform for intelligent agents. *arXiv preprint arXiv:1809.02627* (2020). <https://arxiv.org/pdf/1809.02627.pdf>
- Neil Mendoza. 2019. Robotic voice-activated word-kicking machine. In *ACM SIGGRAPH 2019 Art Gallery*. 1–2.
- Laurent Mignonneau and Christa Sommerer. 2001. Creating artificial life for interactive art and entertainment. *Leonardo* 34, 4 (2001), 303–307.
- Benjamin Péret. 2021. *THE INN OF THE FLYING ARSE*. Retrieved April 4, 2024 from <https://www.surrealism-plays.com/inn-of-the-flying-arse.html>
- Winnie Soon. 2023. *Unerasable Characters Series*. Retrieved April 2, 2024 from <https://calls.ars.electronica.art/2023/prix/winners/7149/>
- Sarah Wasserman. 2020. *The Death of Things: Ephemera and the American Novel*. U of Minnesota Press.
- Wikipedia. 2024. Former Ode on the Red Cliffs. https://en.wikipedia.org/w/index.php?title=Former_Ode_on_the_Red_Cliffs