Between Reality and AI Hallucinations: Reimagining the Sacred Soundscape of Ming Rituals

Ziwei Chi
Tsinghua University
sineewave@gmail.com

Wen Liang
Tsinghua University
liangwen@tsinghua.edu.cn

Abstract

This paper focuses on the soundscape art of the Ming Dynasty's Temple of Heaven sacrificial rituals, employing generative AI hallucinations to reshape spiritual perception and cultural memory. By integrating traditional archival reconstructions including ritual texts, spatial models, and restored music with AI-generated blurred and dynamic audiovisual hallucinations, the work creates a tension between historical rigor and machine ambiguity. A three-screen interactive installation combining traditional soundscape data, AI hallucinations, and a ritual timeline allows audiences to influence the hallucinations in real time, shaping the sacred atmosphere and their individual experience of the ritual soundscape. The project does not pursue historical reproduction but embraces the instability of AI generative systems to explore new possibilities for cultural memory, spiritual experience, and collective imagination.

1 Introduction

Generative AI is often critiqued for producing hallucinations: outputs that deviate from factual accuracy in unpredictable ways. In most engineering contexts, hallucinations are framed as errors to be minimized. This work takes the opposite stance: we embrace hallucination aesthetics as a creative strategy for reimagining intangible cultural heritage. Specifically, we present an interactive installation that reconstructs the sacred soundscape of the Ming Dynasty's celestial sacrifice. It combines historically grounded archival reconstructions with dynamically generated audiovisual hallucinations that respond to audience input in real time.

The Ming-era celestial sacrifice, performed annually at the Temple of Heaven on the winter solstice, was an elaborate ritual involving architecture, choreography, and music. While physical aspects of the ceremony have been documented and partially reconstructed, its perceptual and affective dimensions remain elusive: the way sound shaped spatial experience, the fleeting imagery invoked by chants, the emotional atmosphere of the site. Recovering these experiential layers requires moving beyond static historical reconstruction toward an active negotiation between recorded memory and speculative imagination.

Our installation addresses this gap through a three-layer pipeline:

- Archival Reconstruction – Ritual timeline, spatial layout, and instrument timbres reconstructed from historical documents, archaeological measurements, and iconographic archives.
- Hallucination Generation – Poetic imagery and emotional cues extracted from the archive via NLP drive fine-tuned diffusion models to generate unstable audiovisual outputs, with their sequence structured by historical musical scores.

Corresponding author: Wen Liang (liangwen@tsinghua.edu.cn)

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Creative AI Track.

- Audience Modulation – Ambient sound from visitors modulates the intensity of hallucinations via a real-time mapping function, making the ritual atmosphere responsive to collective presence.

Through this design, the work positions cultural heritage not as a fixed artifact but as a living, negotiated space, continuously reshaped by the interplay of historical fidelity, machine creativity, and collective human presence.

Figure 1: Exhibition Effect of the Soundscape AI Hallucination Installation

2 Background and Related Research: Sacred Soundscapes and AI Hallucinations

Soundscapes are not merely an acoustic domain but a perceptual construct shaped through human experience. Rooted in sound ecology[14] and sound geography[3, 6], this framework emphasizes how listeners and sound together co-create space and meaning, particularly in ritual contexts. In these environments, sound organizes the progress of the ceremony, human behavior, and emotional states, serving as a medium of communication between humans and deities.

Research on sacred soundscapes centers on perception and spiritual experience. Matteo Melioli's concept of "ghost spaces" describes how sound distorts spatial perception, producing complex, non-linear geometrical configurations that extend beyond the physiological boundaries of visual perception. In enclosed spaces, reflected sound projects its source beyond the walls, causing listeners to perceive sound as coming from locations that do not physically exist. This process carves out an expansive imaginative space within the mind, one that sound alone can structure and inhabit[5, 11]. In this sacred context, ethnomusicology places ritual music at the intersection of human behavior, ideology, and sonic materials. Timothy Rice's reconstruction model, drawing on Geertz and Merriam, approaches music through the metaphors of time, space, and musical form, showing how the visible ritual space transforms into a fluid, elastic, and multidimensional domain of imagination[7].

The Ming Dynasty's ritual of celestial sacrifice at the Temple of Heaven exemplifies these qualities. Held at dawn during the winter solstice, the ceremony reduced visual dominance through low lighting while heightening auditory and tactile perception. Chanting, instrument tones, and choreography bound participants, deities, and cosmology into a mechanism of sonic communication and political legitimization. The music and its symbols are inseparable from instrument design and iconography; the "bian qing," a stone instrument often engraved with phoenix patterns, produces a crystalline, ethereal sound that points symbolically to the sanctity of birds. Following Aby Warburg's Mnemosyne Atlas, the ritual's sacred concept operates as a heterotopian matrix of interconnected, nomadic

elements: melodies, timbres, instrument motifs, and lyrics drawn from multiple dynasties, all converging toward a concept of the divine that accumulates and transforms over time[12].

AI hallucinations[8] offer a different angle on this kind of perceptual imagination, one generated not through subjective interpretation but through the abstractions of a machine system. Initially coined in computer vision as a positive super-resolution strategy(e.g., hallucinating facial details) [2], the term was later applied to blurry or fictional outputs in modern generative systems[10]. Recent critiques warn that in factual domains, hallucinated outputs risk being absorbed back into training data, creating a feedback loop that distorts future inferences[4]. In creative practice, however, artists have embraced hallucinations as an aesthetic mode. Memo Akten's *Deep Meditations*[1] moves through latent spaces to produce organic and geological forms in fluid transition, while *ULTRACHUNK* orchestrates a duet between performers and AI avatars, treating ambiguity as a generative condition. Manovich similarly frames AI art as an interface navigating between machine and human perspectives[9].

We consider the Ming Dynasty's ritual soundscape as a negotiated perceptual domain where the historical grounding of archival research and the generative, ambiguous imagination of AI can coexist. We treat hallucinations as a speculative bridge, not an error, using them to reimagine the spiritual perceptions and experiential dimensions of the ritual that remain beyond the reach of written archives.

3 System Design and Artistic Implementation

This research consists of two interdependent components. The first part is a study of traditional soundscapes, where we use the framework of soundscape theory to reconstruct the soundscape of Ming Dynasty celestial rituals based on a large body of historical documents. The data sources include historical records, musical scores, and architectural survey data, which mainly provide macro-level, material aspects of the ritual. The second reshapes this historical foundation through multimodal generative AI, synthesizing textual poetic data, an imagery vocabulary list, and the Chinese Iconography Thesaurus (CIT) [13] to recreate the sensory, speculative, and imaginative experience of ritual "presence."

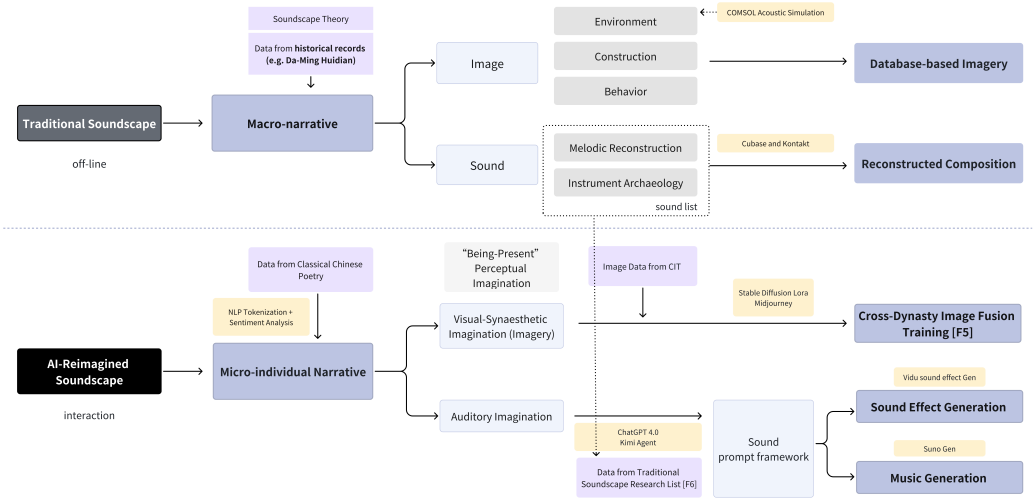

Figure 2: System Process of Traditional Soundscape Research and AI Reconstruction

3.1 Reconstructing the Traditional Soundscape

Soundscape theory focuses on the interaction between people, sound, and the environment. It draws from fields such as landscape studies, psychoacoustics, and architectural acoustics. To reconstruct the sound environment of the Ming Dynasty celestial rituals, we collected and analyzed historical texts, musical scores, and architectural records. Using Comsol Multiphysics®, we modeled the acoustic particles of the environment, combining materials from the Da Ming Hui Dian (Collected Statutes of the Ming) and other imperial archives. This allowed us to construct the spatial layout, music

structure, and specific sound cues for the ritual. We carefully organized the sequence of ritual stages and corresponding musical sections, developing a comprehensive ritual flowchart. For example, in the "Welcoming the Deities" chant, we recreated the scene described in historical texts, where deities descend from the sky in a chariot drawn by dragons and led by a phoenix, accompanied by chimes and bronze bells. This fusion of historical authenticity and ritual atmosphere laid the foundation for AI training and the reconstruction of the ritual.

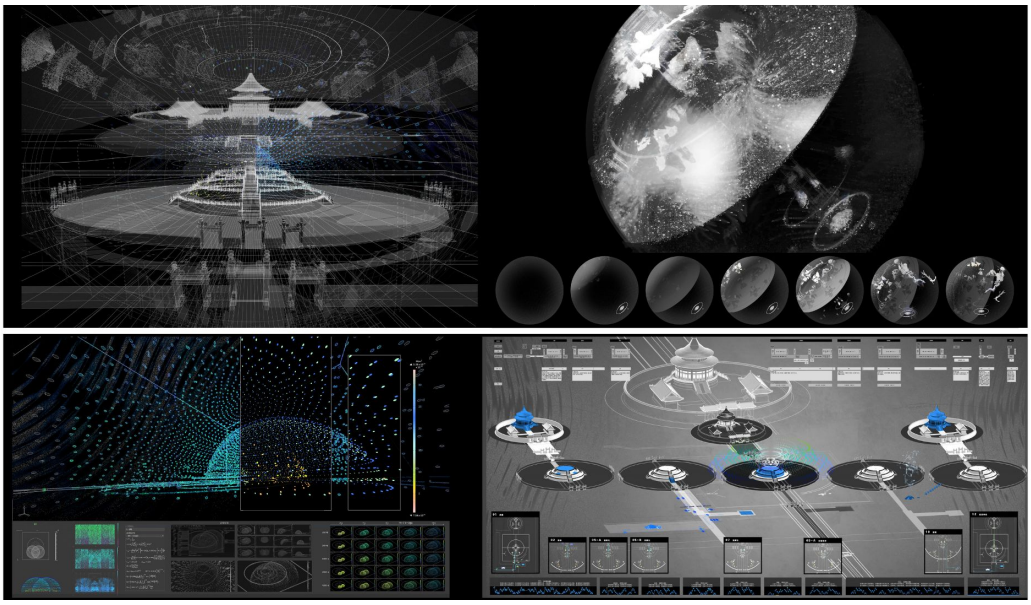

Figure 3: Ritual Soundscape Research in the Ming Dynasty.

3.2 Poetic Data and Imagery List

Traditional soundscape research tends to focus on the macro-level material aspects, often overlooking the participants' sensory and spiritual experiences. To address this, we gathered over a hundred ritual poems written by officials and poets who had participated in the ceremonies. These poems offer detailed descriptions of the sounds, smells, and light of the ritual, adding a spiritual dimension to the concept of "presence." Using NLP technologies such as SnowNLP and deepthulac/LacModel, we performed word segmentation and sentiment analysis, compiling an imagery list that reveals the psychological and spiritual experiences triggered by the ritual. To visualize these poetic images, we extracted nearly a thousand visual archives from the Chinese Iconography Thesaurus (CIT), spanning multiple dynasties, styles, and media. These images were digitized and organized into a non-hierarchical dataset, providing rich visual resources for AI training.

3.3 Generative AI Systems and Hallucination Technology

Building on Deleuze's theory of abstract art, we view the spiritual imagination within sacred rituals as a form of "deliberate abstraction," a conscious choice opposing the figurative. This abstraction is not a mere absence of representation, but a purposeful expression of emotion and experience. AI technology enables this "non-subjective" abstraction. Based on the traditional ritual structure and the poetic imagery list, we incorporated multimodal AI generation technologies to reconstruct the ritual soundscape.

For image generation, we used the NLP-processed imagery list to train AI tools like Stable Diffusion_Lora and Midjourney on images from various dynasties, fusing cultural styles. The goal was to preserve the unique features of each dynasty while merging diverse visual traditions. For example, the image generated for the "Divine Chariot Descending" motif retained the slender, ethereal quality of Han Dynasty stone carvings, alongside the detailed face and hair of a Ming Dynasty dragon.

Figure 4: Image Generation: Cross-dynasty imagistic synthesis + stylization

For sound generation, we designed a multidimensional sound prompt framework using historical musical scores, instruments, and environmental descriptions. This framework combines dynamic music instructions generated by GPT-4, which are then linked to the Vidu sound generator and Suno music synthesis engine. The resulting soundscape includes both instrumental effects and environmental sounds, such as wind and tree rustling, creating an immersive auditory experience. AI hallucinations, often considered "errors," become a productive force here, reflecting the ambiguity and spirituality of the ritual. The blurred, multilayered results break fixed narratives, encouraging the audience to imagine diverse deities and sacred spaces. This aligns with Manovich's view of AI art as a new form lying between "computer systems" and "human aesthetics."

3.4 Interactive Installation: Bridging Archive and Imagination

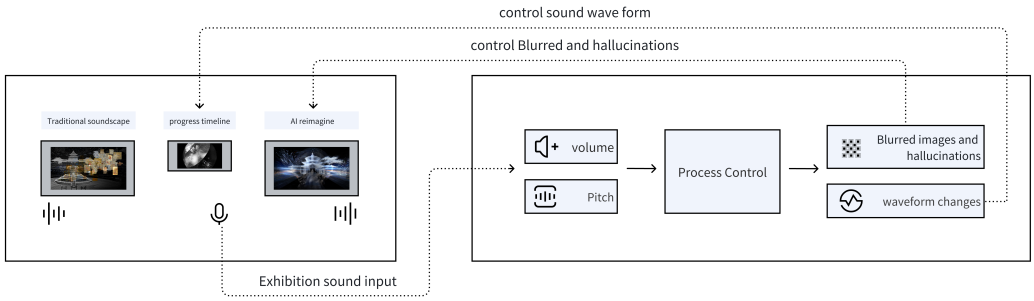

Figure 5: Exhibition Layout and Interaction Logic

The interactive design of the installation features a three-screen layout, fostering a dialogue between tradition and imagination. The left screen displays archival visuals and reconstructed soundscapes based on historical research, showing the material and procedural aspects of the ritual. The right screen presents dynamic hallucination images and sounds generated by AI, reflecting the spiritual ambiguity of the ritual. The central screen serves as a timeline interface, displaying the progression of the ritual from a divine perspective .

Microphones embedded in the space capture real-time audience sounds, analyzing volume and frequency to adjust the timeline waveform and hallucination intensity on the right screen. The more participants there are, and the louder the sound, the richer the hallucinations become, linking collective presence with individual spiritual experiences. This interactive design allows the audience to move between traditional history and generative imagination, enhancing the immersive sense of "presence" and connecting to the spiritual space of the sacred ritual.

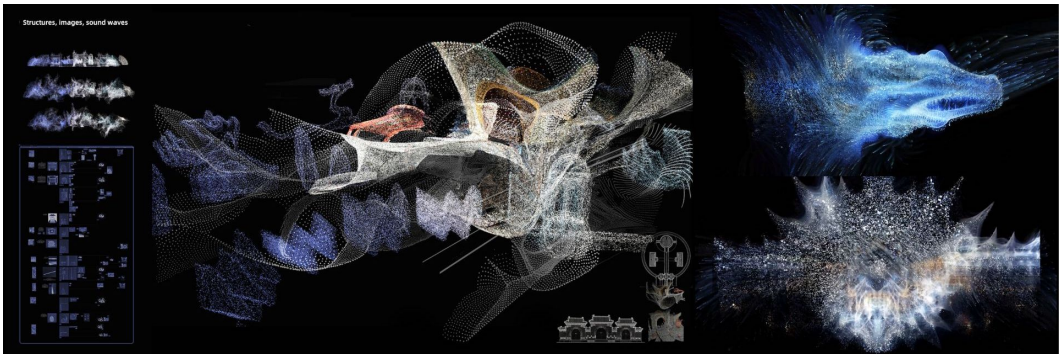

Figure 6: AI-generated audiovisual hallucinations: a particle system merging historical imagery and sonic fluctuations.

4 Discussion: Synthetic Memory, Imagination, and Cultural Futures

This project is grounded in soundscape theory, emphasizing sound as a multidimensional perception of culture and environment. It is not merely a physical phenomenon but a complex construction of social, historical, and spiritual layers. The sound elements and spatial layout in ritual soundscapes jointly shape the participants' sense of presence and spiritual experience, creating a unique contextual atmosphere.

From an iconographic perspective, the project uses cross-dynastic image fusion to reveal the layered superposition of symbols and visual culture, enriching the symbolic meanings of traditional culture. At the same time, individual micro-narratives are embodied in the interaction between soundscapes and AI hallucinations, presenting diverse perceptual viewpoints and mental spaces, breaking the limitations of a singular historical narrative.

AI hallucinations, as a blurry and dynamic generative mechanism, offer an innovative artistic approach for the reconstruction of cultural memory. It is both a technical "flaw" and a catalyst for imagination, fostering the integration of art and technology, and demonstrating the potential of interdisciplinary research methods. Through real-time interaction, the audience becomes a co-creator of cultural representation, enhancing the fluidity of collective memory and the vitality of cultural inheritance.

Looking ahead, with the integration of multimodal data analysis and deep semantic understanding, generative AI is expected to play a more significant role in cultural heritage preservation, digital humanities, and artistic creation, driving a re-examination and expression of historical memory and imaginative space.

5 Conclusion

This work presents a novel approach to reimagining historical soundscapes through the integration of archival research and generative AI hallucinations. By focusing on the Ming Dynasty celestial sacrifice ritual, we reveal how AI's imprecise and ambiguous outputs can become a creative tool to fill perceptual and spiritual gaps left by historical records. Our installation bridges traditional scholarship and immersive interactive experience, inviting audiences to engage with both the tangible and imagined layers of ritual soundscapes. This synthesis opens new pathways for cultural memory reconstruction, collective imagination, and interdisciplinary exploration.

Future work could deepen the integration of AI hallucinations with multimodal sensory inputs and expand the framework to other cultural and historical contexts. We anticipate that such synthetic memories will enrich how we perceive, preserve, and reinterpret intangible heritage in an increasingly digital world.

References

[1] Memo Akten, Rebecca Fiebrink, and Mick Grierson. Deep meditations: Controlled navigation of latent space. *arXiv preprint arXiv:2003.00910*, 2020.

[2] Simon Baker and Takeo Kanade. Hallucinating faces. In *Proceedings Fourth IEEE international conference on automatic face and gesture recognition (Cat. No. PR00580)*, pages 83–88. IEEE, 2000.

[3] Philip Dunbavin. Iso/ts 12913-2: 2018—soundscape—part 2: Data collection and reporting requirements—what's it all about. *Acoust. Bull*, 43:55–57, 2018.

[4] Ziv Epstein, Aaron Hertzmann, Investigators of Human Creativity, Memo Akten, Hany Farid, Jessica Fjeld, Morgan R Frank, Matthew Groh, Laura Herman, Neil Leach, et al. Art and the science of generative ai. *Science*, 380(6650):1110–1111, 2023.

[5] Yixiao Fu, Daragh Byrne, and Lawrence Shea. Evoking the post-industrial landscape memories through spectrality and mixed reality soundscapes. In *Proceedings of the 13th Conference on Creativity and Cognition*, pages 1–6, 2021.

[6] Michael Gallagher and Jonathan Prior. Sonic geographies: Exploring phonographic methods. *Progress in human geography*, 38(2):267–284, 2014.

[7] Clifford Geertz. *The interpretation of cultures*. Basic books, 2017.

[8] Negar Maleki, Balaji Padmanabhan, and Kaushik Dutta. Ai hallucinations: A misnomer worth clarifying. In *2024 IEEE Conference on Artificial Intelligence (CAI)*, pages 133–138, 2024.

[9] Lev Manovich. Defining ai arts: Three proposals. *AI and dialog of cultures" exhibition catalog. Saint-Petersburg: Hermitage Museum*, 2019.

[10] Søren Dinesen Østergaard and Kristoffer Laigaard Nielbo. False responses from artificial intelligence models are not hallucinations, 2023.

[11] Colin Ripley, Marco Polo, and Arthur Wrigglesworth. *In the place of sound: architecture, music, acoustics*. Cambridge Scholars Publishing Newcastle, 2007.

[12] Cristina Tartás Ruiz and Rafael Guridi García. Cartografías de la memoria. aby warburg y el atlas mnemosyne. *EGA. Revista de expresión gráfica arquitectónica*, 18(21):226–235, 2013.

[13] Chinese Iconography Thesaurus. Chinese iconography thesaurus, 2025. Accessed February 2025.

[14] Barry Truax. Acoustic ecology and the world soundscape project. *Sound, media, ecology*, pages 21–44, 2019.

A Supplementary Material

A.1 Previous Exhibition: Original Version

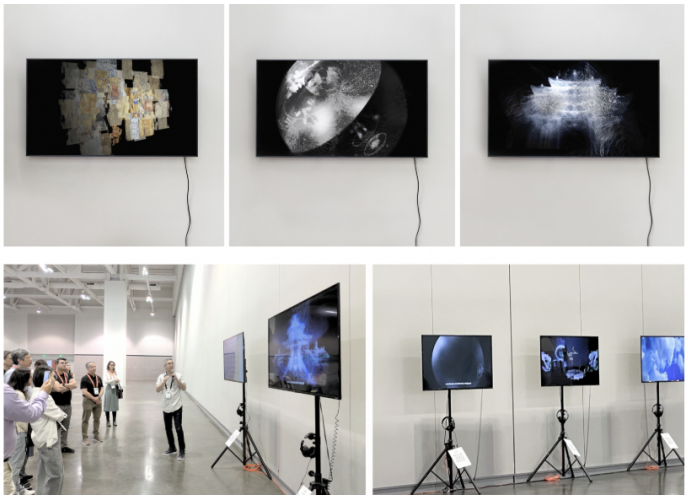

Figure 7: Exhibited at CVPR 2025 Art Gallery, Music City Center, Nashville, TN, USA

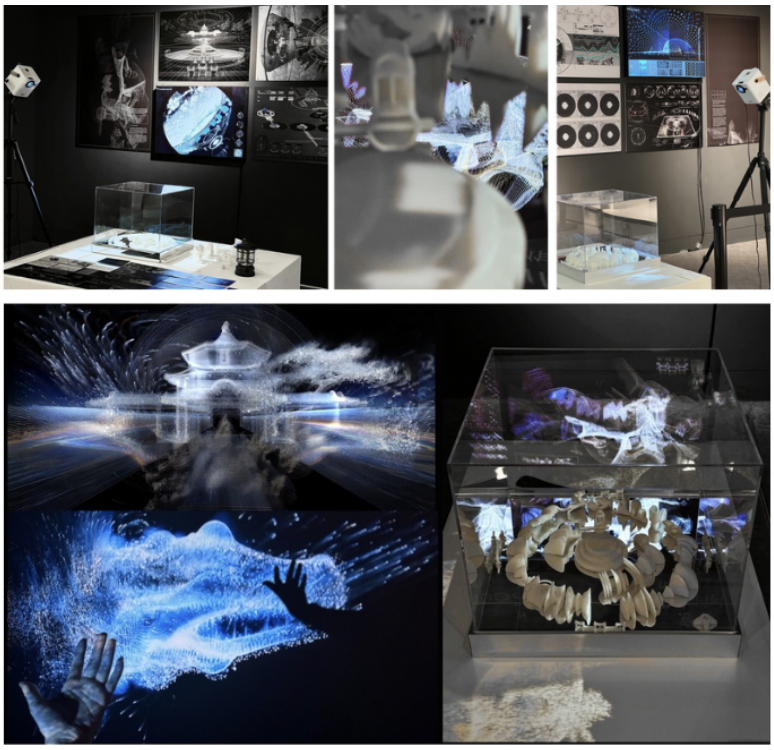

Figure 8: Exhibited at Tsinghua University, Beijing, China

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction have stated the contribution: a three-layer pipeline and an interactive installation that fuses archival reconstruction with AI hallucinations.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The paper states it does not aim for faithful historical reproduction and embraces AI instability. We left the generalization to other contexts for future work.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [NA]

Justification: This work presents a system and an interactive installation, and does not cover formal theorems or proofs.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [NA]

   Justification: As the aim of our work is to present an interactive installation and qualitative system description, quantitative experiments are not applied.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [No]

Justification: While the pipeline, tools, and data sources are described, datasets are not released at submission time due to license issues.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [NA]

   Justification: No training/testing experiments or benchmarks are reported that would require train/test splits or hyperparameter details, as the aim of our work is to present an interactive installation and qualitative system description.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [NA]

   Justification: The paper does not include statistical experiments; no error bars or significance tests are applicable.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.

- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [NA]

   Justification: As the aim of our work is to present an interactive installation and qualitative system description, no computational experiments are reported that require per-run compute disclosure.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: This work adheres to the NeurIPS Code of Ethics: no human-subject data collection or PII; ambient sounds are used only for real-time modulation in the installation.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: The paper discusses positive cultural-heritage impacts in the Discussion, and covers concerns around generative AI feedback contamination in the background.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [NA]

    Justification: No high-risk models or scraped datasets are released; the work uses existing generation tools within an installation context.

    Guidelines:

    - The answer NA means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: The models used in the paper are open-source models, and the data are open-source datasets (under their respective licenses). All other assets, such as images and architectural modeling, are created by the authors themselves.

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.

- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [NA]

    Justification: No new datasets, models, or code are released alongside the paper.

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification: The work involves audience interaction in a gallery setting but does not conduct crowdsourcing or human-subject studies for research analysis.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: No crowdsourcing or human-subject research is reported that would require IRB review.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

    Answer: [Yes]

    Justification: GPT-4 are used to generate dynamic music instructions as a non-standard component of the audiovisual pipeline (Sec. 3.3).

    Guidelines:

    - The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.