



Article

An Exploratory Analytical Framework for Artificial Intelligence-Driven Interactive Art Installations

Sihwa Park ^{1,*}  and Je-ho Oh ^{2,*}

¹ Computational Arts, School of the Arts, Media, Performance & Design, York University, Toronto, ON M3J 1P3, Canada

² Department of Metaverse & Game, Soonchunhyang University, Asan 31538, Republic of Korea

* Correspondence: shpark@yorku.ca (S.P.); ohjeho@sch.ac.kr (J.-h.O.)

Abstract

Interactive art is a form in which aesthetic content emerges through audience participation within technology-based installations designed for interactive engagement. Recent developments in AI have increasingly shaped this field, influencing audience participation, the dynamic interplay between installations and algorithms, and the formation of aesthetic experience. However, relatively little research has examined how AI affects these design dimensions in interactive art. To address this gap, this study proposes the exploratory 4A framework, comprising audience participation, art object embodiment, AI inquiry and response, and aesthetic experience. Rather than advancing a wholly new theory of interaction, the framework extends earlier models by treating AI as an analytically distinct entity to examine the technical and experiential conditions specific to AI-driven artworks. Drawing on publicly available documentation, we analyzed three artworks presented at the CVPR 2024 AI Art Gallery as illustrative cases to demonstrate the framework's application. The analyses elaborate interpretive design propositions concerning AI temporality, modality conversion, and AI-mediated aesthetic experience. These propositions suggest potential approaches to designing AI-driven interactive installations, while the framework provides a structured basis for further analysis. The study's limited case selection and documentary basis highlight the need for broader application, framework refinement, and empirical investigation of audience experience.

Keywords: artificial intelligence; interactive art; interaction design; audience participation; aesthetic experience



Academic Editor: Stephan Schlögl

Received: 3 August 2026

Revised: 25 September 2026

Accepted: 29 September 2026

Published: 1 October 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

1.1. Definition and Aesthetic Effects of Interactive Artworks

Interactive art is a genre in which artistic content is generated in real time through the direct participation of an audience, mediated by algorithms and digital technologies within an interactive installation [1–3]. Because audience involvement is foundational, these works are characterized as participatory art that continuously evolves. Unlike traditional artworks, which primarily emphasize appreciation and interpretation, contemporary interactive art elevates real-time participation to a central element in the creation of artistic content and subsequent aesthetic experience. This shift signifies the emergence of a new form of aesthetic value.

Aesthetic interaction instantiates the human effort to find meaning when encountering art; a viewer's interpretive ventures are discursive accomplishments that integrate the aesthetic dimension of life into the broader human experience [4].

Interactive art installations have continually evolved through the application of digital technologies. Consequently, prior research has explored the design and function of installations that enable interactivity [5–9]. By combining various technologies, these interactive installations can reflect audience actions and, in doing so, generate novel aesthetic experiences.

The implementation of interactive art has spurred continued research into new technological applications, the nature of interaction, and the temporal and spatial dimensions of content experience [10–14].

A fundamental relationship between the audience and the installation has been established in this field since the work of Fels [15]. This foundational structure has since evolved into increasingly complex and participatory forms, as documented by Edmonds [16]. The application of diverse technologies has enabled new participation modes and facilitated the integration of different content types within the structure of interactive art. Consequently, audience participation, technological mediation, and artistic representation have emerged as central research themes [5–17].

In particular, recent advancements in artificial intelligence (AI) technologies have introduced significant changes and expanded the possibilities for production techniques, modes of audience participation, and the nature of the aesthetic experience in interactive art. From the perspective of installation design, integrating various AI algorithms has enabled new forms of engagement that generate aesthetic experiences distinct from those possible with traditional interactive art.

1.2. Integrating AI into Artistic Practice

AI technologies, propelled by an explosive increase in data, enhanced computing power, and breakthrough model architectures, are influencing virtually every aspect of modern life. At its core, AI aims to develop intelligent agents (systems) capable of reasoning, evolving, and functioning autonomously through learning. This integration of AI has catalyzed new forms of artistic transformation in interactive artworks [18–24], and academic research is actively investigating the influence of AI in the arts, including its implications for ethics and creativity [25,26].

Even before the advent of generative AI, machine learning (ML) methods—including supervised learning, unsupervised learning, and reinforcement learning [27]—were employed in the arts. These methods were particularly used to produce personalized content, generate creative output, and support artistic ideation [28]. The socio-technical conditions established for producing art-based ML prototypes have paved the way for an interactive understanding of ML, one that embraces the differences between design and engineering practices through diffractive practice [29].

Interactive ML has also been used to design and build new gestural controls, allowing artists and dancers to define embodied interactions through improvisation and evaluate models in real time by experimenting with controllers [30]. The application of deep learning in art is enhancing artistic creativity, influencing both ideation and visualization by enabling new modes of expression [31,32].

More recently, the emergence of generative AI—driven by models such as variational autoencoders (VAEs), generative adversarial networks, diffusion models, and transformer-based large language models (LLMs)—has begun to challenge traditional notions of human creativity. Generative AI has opened new opportunities for audience engagement

through image generation, transforming how audiences create and experience visual content [33].

Furthermore, co-creation between humans and AI is provoking ongoing discourse regarding the evolving roles of human creativity and technological agency [26,34,35]. As the scope of AI-driven art production and exhibition expands, it prompts sustained academic exploration of its aesthetic significance [36,37]. Consequently, transformative changes are occurring across the entire artistic process, from data collection and AI system design to technological procedures, aesthetic representation, and modes of audience participation.

1.3. Study Aims

This study aims to advance the analysis of AI-driven interactive art by examining recent interactive installations that incorporate AI. Although prior research has established important frameworks for traditional interactive art, relatively few studies have addressed how AI specifically reshapes interaction design, system behavior, and aesthetic experience in interactive installations. This gap persists owing to the limited capacity of traditional interactive art frameworks to address AI-specific design implications and the lack of established methodologies for examining how AI integration influences audience participation and aesthetic experience.

To address this gap, this study proposes an exploratory analytical framework for AI-integrated interactive art, informed by initial artwork reviews and engagement with existing models of interactive art. Rather than advancing a wholly new theory of interaction, the framework organizes the technical and possible experiential conditions of AI-driven artworks. Its development and application were interrelated. Although the final framework is presented before the detailed case analyses for clarity, this order of presentation does not describe a strictly linear order of discovery. The research developed through the following overlapping stages:

1. Identify interactive AI artworks among works featured in the AI Art Gallery of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2024.
2. Conduct initial reviews of selected artworks using a common exploratory template.
3. Develop and refine a more concrete analytical framework in relation to these reviews and the relevant literature on interactive art, AI, and aesthetic experience.
4. Analyze three cases in detail using the framework to demonstrate its application.
5. Elaborate design-oriented interpretations concerning interaction temporality, modality conversion, and AI-mediated aesthetic experience.

1.4. Scope and Limitations

The scope of this study is as follows:

- (1) The analysis is based on official artwork descriptions, related publications, and publicly available video documentation, not on primary audience interviews, surveys, or behavioral measurements. It is a documentary and interpretive analysis. Descriptions of system operations are distinguished from researchers' interpretations of possible experiences. The case analyses demonstrate the exploratory framework's application; they do not independently validate it or establish actual audience responses.
- (2) Ethical issues related to the use of AI, including consent, privacy, and bias, are highly significant in contemporary AI art. While these concerns are acknowledged and briefly discussed in this study, they are not examined as a primary analytical focus.
- (3) Although theoretical questions concerning artist agency, authorship, and aesthetic value are central to ongoing discussions on AI-based interactive art, they are not the main focus of this study. The study is intended to advance an analytical framework rather than to propose new theories regarding these issues.

2. Literature Review

To analyze AI-integrated interactive art installations, a literature review was conducted across two related strands of literature: interactive art more broadly and AI-integrated interactive art installations specifically. The review aimed to examine the interaction structures of interactive art installations, explore the roles of AI technologies within these structures, and identify characteristics associated with AI integration. It also considered how AI reconfigures relationships among audiences, artworks, and computational systems, and how these relationships shape aesthetic experience. The review adopted a narrative approach and conceptual synthesis rather than a systematic review designed to provide an exhaustive synthesis of evidence.

Relevant literature was identified through searches in Google Scholar, Scopus, IEEE Xplore, and the ACM Digital Library, supplemented by citation tracking of key publications. Search terms included combinations of “interactive art,” “AI interactive art,” “artificial intelligence interactive installation,” “generative AI art,” “human–AI interaction,” “audience participation,” “interactive aesthetic experience,” “computational creativity,” and “AI-mediated interaction.” Reference lists of key publications were also examined to identify foundational studies and established models relevant to the conceptual analysis of AI-mediated interactive systems.

The review first examines established models and conceptual accounts of interactive art installations, including input–output processes, audience participation, embodiment, and artwork–audience relationships. It then considers studies of AI-based interactive installations, examining how AI technologies are incorporated into artistic practices and contribute to co-creation, audience engagement, and aesthetic experience. Together, these strands provide the conceptual foundations and examples that inform this study’s proposed analytical framework for examining the relationships between AI technologies, interaction design, and the artistic and conceptual implications of interactive installations.

2.1. Research on Interactive Art Installations

Over the past decade, interactive art installations have become a significant research topic at the intersection of art, technology, and human–computer interaction (HCI). Such installations employ various technologies to capture and respond to audience activities, generating new aesthetic experiences through dynamic interaction.

Saltz [37] conceptualized a process in which a sensing or input device translates aspects of a person’s behavior into a digital form understood by a computer; the computer then outputs data that are systematically related to the input, which are then translated back into real-world phenomena. Graham [38] explored the development of interactive artworks enhanced by collective usage, applying a metaphor of “conversation/host” to the creation of the artwork; Graham also conducted further case studies and developed a taxonomy to illustrate differences in artwork–audience and audience–audience relationships. Fels [15] discussed three artworks—Iamascope, Video Cubism, and The Forklift Ballet—that combine technology and art to illustrate issues of intimacy and embodiment. Edmonds et al. [39] discussed the role of interaction in art systems and new methods for building them, while Edmonds [16] later developed a framework defining audience participation through sequential phases: adaptation, learning, anticipation, and deeper understanding. Xiaobo and Yuelin [17] investigated a basic interaction framework integrating time, space, and information within interactive systems to enable varied artistic experiences. Ahmed [5] analyzed how interaction and interactivity are defined in different fields, examining their relevance and applicability in the context of digital interactive art.

These interactive art installations incorporated various technologies to reflect audience actions, creating new aesthetic experiences through dynamic interactions.

2.2. Research on AI-Based Interactive Art Installations

AI has been increasingly employed in the field of arts, functioning both as a computational instrument and an autonomous system capable of producing aesthetic outcomes. AI technologies have also been incorporated into installation works that create interactive and adaptive experiences. Advances in AI have further facilitated interactions between humans and intelligent systems, establishing new paradigms for creative collaboration and audience engagement.

The application of AI in this context can be categorized into broad tiers, including AI, ML, deep learning (DL), and generative AI approaches. For instance, Chen et al. [40] built a new interactive art creation system with AI as the core medium, using a methodological approach of “cognition of human–computer symbiosis, innovation supported by intelligent technology, collaboration among creative subjects, and constraints on creative behavior,” where artists, robots, and the audience are co-creation subjects. Xie et al. [41] examined the integration of VAEs and particle swarm optimization in generative and interactive digital installations, revealing that this hybrid model enhances user engagement and aesthetic satisfaction. Canet Sola and Guljajeva [42] proposed the Dream Painter, an interactive robotic art installation that turns spoken dreams into a collective painting. They explored the interactive potential of AI and robotics, sparking a broader discussion on DL applications. In their other work, Visions of Destruction [43], they used artistic research methods to communicate a focus on the Anthropocene through audience interaction and generative AI [18]. Sun et al. [22] presented AI Nüshu, an emerging language system inspired by Nüshu, a unique language used exclusively by ancient Chinese women. Their work highlighted the role of AI in preserving cultural heritage and redefining human–machine dynamics in a language-making interactive artwork. Zhang et al. [24] described the conceptual background, AI system design, and visualization strategies of ReCollection, an interactive AI art installation. The artwork assembles collective visual memories from the language input of participants, blurring the boundaries between remembrance and imagination.

These works have employed generative AI technologies to create artistic content such as images, sound, and video, thereby offering new functionalities within interactive installations and expanding the aesthetic experiences of the audience.

While research on interactive art installations has received considerable attention, studies that focus specifically on the underlying frameworks of AI-based interactive artworks to better understand the impact of AI technology on interaction design and aesthetic effects remain limited.

Moreover, prior research has largely focused on the technical implementation of AI-based artworks, rather than offering a comprehensive analytical perspective. To address these gaps, this study proposes an exploratory analytical framework and case-based approach to examine the artistic and conceptual implications of AI-driven interactive installations.

3. Research Framework

This section reviews existing research frameworks for interactive art and presents a new framework for analyzing interactive installation art that incorporates AI technologies.

3.1. Research Framework Development

The framework originated in exploratory initial reviews of selected interactive AI artworks presented at CVPR 2024 using a common template: input, output, temporality, AI technologies, and key points. These were researcher-devised descriptive fields and an open-ended analytical note section, rather than a validated or exhaustive taxonomy. They provided an initial structure for examining how audience contributions enter an instal-

lation, how computational responses are produced and presented, and how interaction unfolds over time.

Input and output were relevant to the study's focus on the relationship between audience participation and the installation's response, consistent with the interaction process discussed by Saltz [37]. Temporality directed attention to timing and sequencing rather than input and output modalities alone, while AI technologies identified the models and functions through which AI participated in the system. The key-points section recorded technical observations, interaction-related observations, uncertainties, and aesthetic interpretations; it was not treated as an additional dimension of the artwork.

Framework development combined case-informed interpretation with literature-based conceptual synthesis. The initial reviews provided descriptive observations and provisional analytical concerns, which were subsequently organized and reconsidered in relation to existing models of interactive art and literature on aesthetic experience. A more concrete framework was initially conceived as the principal research output and was subsequently used and refined through detailed case analysis. This relationship between development and application has precedents in HCI: Hornecker and Buur [44] developed their framework through system studies and literature synthesis and subsequently used related cases to demonstrate its application, while Benford et al. [45] refined an emerging conceptual framework through retrospective examination of published studies. In the present study, the detailed cases similarly serve as worked applications of an evolving framework, rather than as an independent validation dataset.

Previous studies have proposed fundamental models of interaction between humans and objects. Fels [15] categorized four types of relationships based on the extent to which an object is embodied within a person or the person within the object:

1. The person engages in dialogue with the object,
2. the person embodies the object,
3. the object communicates with the person, and
4. the object embodies the person.

This foundational model provided a methodology for categorizing communication between audience, object, and installation.

Edmonds et al. [39] delineated the relational structure between artwork, artist, viewer, and environment in interactive art using four categories: static, dynamic-passive, dynamic-interactive, and dynamic-interactive (varying). These categories reflect the extent and nature of the interaction achieved.

Xiaobo and Yuelin [17] structured their "Interaction Aesthetics Framework" into three components: interaction (classified into time, information, and space); interactive system (classified into function, form, and structure); and aesthetics (classified into visual, auditory, and tactile domains). This framework made it possible to distinguish characteristics of audience engagement and systematically articulating aesthetic experiences shaped by the audience.

Ahmed [5] focused on the communicative relationship between humans and media within the context of interactive art, proposing a five-stage classification model:

1. Human–Human Interaction,
2. Human–Media Interaction,
3. Device-Mediated Communication,
4. Human–Machine Interaction, and
5. Triggers for Change.

This model enabled a nuanced understanding of the interrelationships between audience, installation, environment, and aesthetics.

Guljajeva and Canet Sola [42] introduced a triangular framework of interactive art comprising author, artwork, and audience. In this model, feedback between author and artwork is generated through code and variables, whereas feedback between author and audience emerges through parameters, system stress, and user interaction. They proposed that the aesthetics of interactive artworks are shaped by the reciprocal exchange of input and output between audience and artwork. Fundamentally, their study conceptualized interactive art as a structure wherein installations—including digital devices designed in alignment with the intent of the artist—generate artistic content through audience interaction, which is then realized and experienced through participation.

The fundamental models discussed above are summarized and compared in Table 1. Within these models, AI may be understood either as part of the art object, that is, as a component of the installation system, or as a software-based agent embedded within it. However, this perspective becomes insufficient for AI-driven interactive artworks, where AI is not merely embedded in the installation system but actively mediates interaction, generation, and interpretation. The incorporation of AI has expanded interaction modalities, introduced new thematic and conceptual concerns, and enabled artworks to operate with partial autonomy while forming dynamic relationships with viewers [46]. Therefore, AI should not be reduced to a generative algorithm or technical subsystem within the art object. It must be treated as a distinct analytical entity, since its inferential processes, generative uncertainty, and relative autonomy exert a direct influence on how interaction unfolds and how aesthetic experience is produced.

Table 1. Summary of fundamental models.

Categories	Fels [15]	Edmonds et al. [39]	Xiaobo & Yuelin [17]	Ahmed [5]	Canet Sola & Guljajeva [42]
Human	Person	Viewer, artist (human agent)	Audience	Artist, audience	Author, audience
Objects	Object	Art object, environment, software agent	Interactive system	Artwork, environment	Artwork
Relationships	Engagement, embodiment, communication	Interaction (Static, dynamic-passive, dynamic-interactive, dynamic-interactive (varying))	Interaction, aesthetics	Interaction, communication	Feedback

To address this gap, we propose the exploratory 4A framework for the analysis of AI-based interactive art. The framework consists of three key entities: audience, art object, and AI, and four key design elements: participation, embodiment, AI inquiry and response, and aesthetic experience (see Figure 1). Its proposed contribution lies in analytically distinguishing AI's inferential and generative operations from the installation's material presentation, while examining how these operations relate to audience participation and potential aesthetic experience. Compared with the models reviewed above, the framework treats AI as an analytically distinct entity rather than subsuming it within the art object or installation system. Table 2 summarizes the framework, which is offered for structured analysis and further development rather than as an empirically validated model.

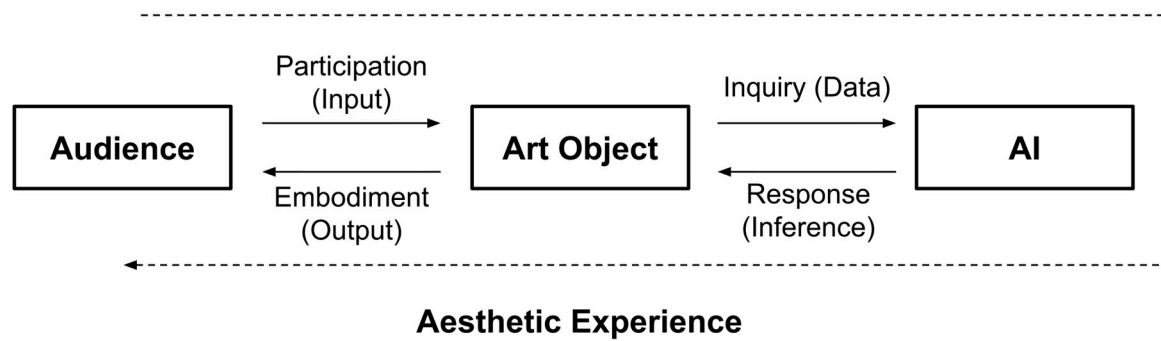


Figure 1. Proposed 4A Framework. The dotted arrow indicates that aesthetic experience may arise across all stages of interaction among the audience, art object, and AI, rather than only from the final output.

Table 2. Core Elements of the 4A Framework.

Key Design Elements	Analysis Elements	Description
Audience participation	Input	Audience participation is closely tied to the system's input. This element is categorized by the format and modality of the input data, including participant choices that influence the artwork's overall shaping process.
Art object embodiment	Output	Final format and modality of the system's output.
AI inquiry and response	AI technologies	AI models or methods employed.
	Data processing	Data pre-/post-processing for inquiry and inference.
	Temporality	Timing and sequencing of audience input, AI processing, and output updates; continuity and delay are distinguished.
Aesthetic experience	Sensory, perceptual, and cognitive experience	Interpretation of possible sensory, perceptual, and cognitive experiences in relation to documented interaction and presentation. Not a measurement of actual audience response.

3.2. Elements of the 4A Framework

Each element within the proposed 4A framework is described as follows:

1. Audience Participation

Interaction between the audience and the artwork is initiated and mediated by the input data provided by participants. This framework analyzes the structural characteristics of input data (format) and sensory channels (modality) to develop a deeper understanding of the interaction process. This also includes choices that may influence the overall shaping process of the artwork system.

2. Art object embodiment

The framework investigates how outputs are generated and presented within the installation system. Similarly to the analysis of inputs, this involves a detailed examination of the format and modality of output data, including sensory channels (e.g., visual, auditory) and structural forms through which the results are materialized or experienced.

3. AI inquiry and response

- (a) **AI Technologies:** The framework identifies the specific AI models, algorithms, and tools integrated into the interactive installation. It distinguishes the func-

tional types of AI technologies (e.g., image generation, object recognition, speech processing) and, where applicable, subtypes (e.g., image-to-image or text-to-image generation). These can be realized via publicly available pre-trained models (e.g., Stable Diffusion [47], BLIP-2 [48]) or artist-developed, fine-tuned models designed to meet particular conceptual and aesthetic intentions.

- (b) **Data Processing:** The framework examines how audience input data are preprocessed before being fed to the AI model. It also investigates whether supplementary data are incorporated and whether the output of the AI model is used directly to materialize the artwork or requires additional post-processing. Furthermore, the framework considers whether accumulated data from previous interactions or environmental sensing are integrated with new inputs during the ongoing interaction.
- (c) **Temporality:** The framework addresses the temporal characteristics of interaction within AI-driven installations. While temporality is also shaped by the broader processes of audience engagement and artwork manifestation, the framework treats AI as the primary factor and therefore locates temporality within AI inquiry and response rather than the art object itself. This also includes the time required to process the audience input and produce output, and the temporal design of interaction modes and pacing of the manifestation of the artwork. Interaction processes are categorized as procedural, delayed, near-immediate, and real-time, depending on system responsiveness and the continuity of user engagement. However, these categories are illustrative temporal arrangements, not mutually exclusive latency classes. No quantitative latency thresholds are inferred without supporting evidence.

4. Aesthetic Experience

Aesthetic experience is inherently subjective and may vary considerably in interactive art through different forms of audience participation. Previous studies describe aesthetic experience as involving interconnected sensory, perceptual, cognitive, and affective processes through which audiences perceive stimuli and construct meaning [49–57]. These dimensions provide a conceptual basis for considering potential aesthetic experience in interactive artworks without assuming uniform audience responses.

In AI-integrated installations, these processes may be further shaped by technological mediation. AI systems generate and transform content through datasets, model architectures, prompts, and probabilistic inference, potentially introducing uncertainty, biases, and unexpected associations [25,46,58]. Such characteristics may influence both the sensory outputs of an artwork and the conditions through which they are perceptually and cognitively interpreted, potentially affording experiences such as surprise, curiosity, confusion, or interpretive ambiguity.

Accordingly, this study relates documented system characteristics, including generative uncertainty and multimodal interaction, to potential aesthetic experience. The analysis distinguishes documented technological, interactional, and contextual information from the authors' interpretations. Because no primary audience data were collected, affective and experiential qualities are treated as potential experiential affordances rather than empirically validated audience responses.

4. Case Studies

This section distinguishes the broader candidate inventory, the six exploratory reviews that informed framework development, and the three detailed illustrative analyses.

4.1. Selection and Role of the Detailed Cases

The broader candidate inventory comprised 13 works identified as interactive installations incorporating AI within the 116 entries in the CVPR 2024 AI Art Gallery, based on the online descriptions and available video documentation. Table A1 in Appendix A lists this inventory. The gallery context identifies the source of the works; it should not be read as evidence that all 13 received equivalent detailed analysis.

Six works received initial reviews using the shared exploratory template: Unreal Pareidolia -shadows-, ReCollection, Beyond Characters: The Unseen Labyrinth, Visions of Destruction, AI Nüshu, and The Endless Collaborative Book: Wushu. This exploratory subset was not assembled through a predefined systematic sampling protocol. The initial reviews, including those of works that were not retained for detailed presentation, informed framework development. Appendix B summarizes the six exploratory reviews that informed framework development, retaining the original template and recorded uncertainties.

To keep the manuscript focused, we limited the detailed case presentations to three works. The principal selection consideration was that the exploratory reviews revealed contrasting strategies for organizing the temporal relationship among audience input, AI processing, and visible output. We characterized these as procedural or delayed interaction, pseudo-real-time interaction, and continuous or real-time interaction. These labels refer primarily to the temporal organization of interaction rather than to mutually exclusive latency ranges:

1. Unreal Pareidolia -shadows- [59,60] illustrates procedural, staged interaction: participants explicitly confirm a physical arrangement, after which caption generation and image generation occur sequentially before the final output is presented.
2. ReCollection [24,61,62] illustrates pseudo-real-time interaction: speech capture, transcription, narrative processing, and image generation form a sequential pipeline, with an approximately 20-second generation delay. During this interval, visual memories from the previous input fade and dissolve into the newly generated imagery, bridging the delay through a continuous visual transition.
3. Visions of Destruction [18,43,63] illustrates continuous interaction: real-time gaze sensing is coupled with near-real-time AI inpainting and continuously evolving imagery. Thus, sensing, processing, and visual updating operate at different temporal rates within an ongoing interaction loop.

The selection was thus temporality-guided but was not governed by a formal sampling codebook or an exhaustive comparison across all 13 works. The three cases provide illustrative contrasts rather than an exhaustive temporal taxonomy or representative coverage of interactive AI art. Figure 2 summarizes the relationship among the exploratory reviews, framework development, and detailed case analyses, while Table 3 identifies the three selected cases.

Table 3. Artworks Retained for Detailed, Temporality-Guided Illustration.

Artist(s)	Artwork Title	Year
Scott Allen	Unreal Pareidolia -shadows-	2023
Weidi Zhang and Rodger Luo	ReCollection	2022–2024
Varvara Guljajeva and Mar Canet	Visions of Destruction	2023

Limiting the detailed presentation to three works was a decision about the focus of the paper, not a demonstration of sample sufficiency or analytic generalization. The cases are situated within qualitative, design-oriented inquiry at the intersection of art and HCI [64]. In this study, however, their role is to demonstrate an evolving analytical framework and

elaborate design propositions, not independently validate its dimensions or establish the prevalence of the proposed patterns.

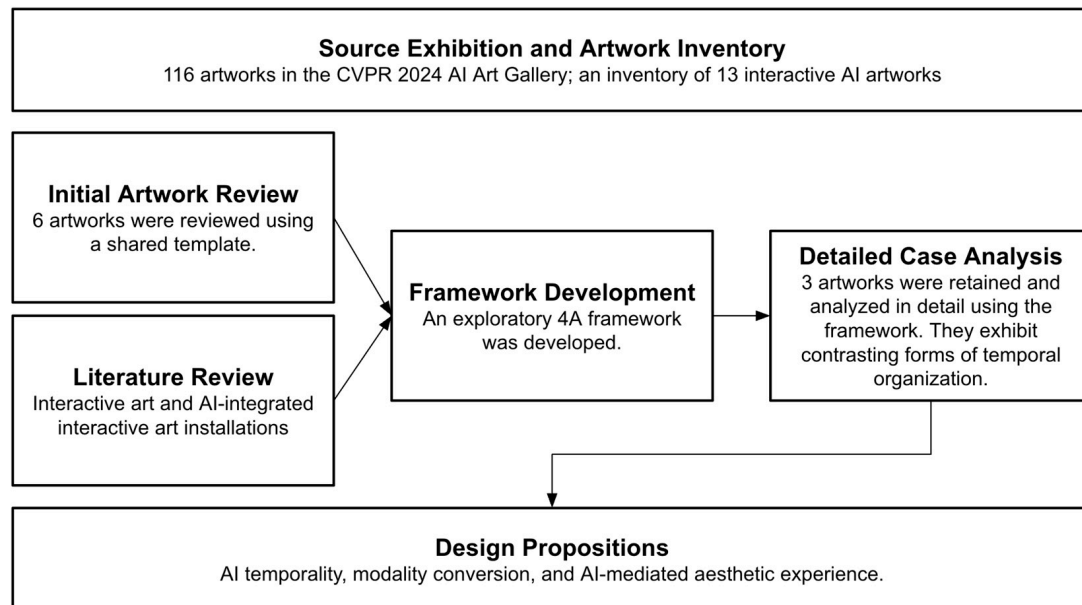


Figure 2. Overview of exploratory framework development and case analysis. The inventory provides context, not evidence of an independently coded 13-case selection protocol.

4.2. Analytical Approach

The detailed analyses draw on official artwork descriptions, related publications, and publicly available video documentation. They are organized by the evolving 4A framework. The shared initial review template provided descriptive structure, but the study did not employ a predefined formal coding scheme or an inter-rater assessment. The analytical questions below clarify the framework's application; they should not be read as a retrospectively asserted independent coding protocol.

For each detailed case, the analysis addresses the following elements:

1. Audience participation: Describe the participant's actions and contributions, the captured input, and the choices that influence the installation. Distinguish an audience action from the data representation supplied to an AI model.
2. Art object embodiment: Describe the format and modality of the output and how it is presented. Distinguish generated content from animation, projection, or other presentation processes.
3. AI inquiry and response: Examine AI functions, data transformations, and temporal organization:
 - a. AI technologies: Identify reported models and functions. Do not infer an architecture, training dataset, or fine-tuning procedure from the artwork's theme or a model name.
 - b. Data processing: Trace the described relationships among captured input, pre-processing, model conditioning, inference output, and post-processing. Record gaps rather than filling in undocumented implementation details.
 - c. Temporality: Describe initiation, input continuity, processing sequence, update timing, and output continuity separately. Distinguish source-reported responsiveness from measured latency and from inferred audience perception.
4. Aesthetic experience: Relate the described interaction and sensory presentation to possible perceptual and cognitive interpretations. Attribute artistic intentions to the

source and label researchers' interpretations as potential experiential affordances, not observed audience effects.

Across these elements, reported system operations, artist-stated intentions, and researchers' interpretations are distinguished in the wording of the analysis. The subsequent discussion elaborates temporal concerns that informed selection and develops additional interpretations concerning modality conversion and AI-mediated aesthetics. Proposed applications beyond the analyzed works are design propositions, not implementations or effects demonstrated by these cases.

Table 4 summarizes the analytical questions and evidentiary boundaries in this revised presentation. It provides guidance for reading and applying the exploratory framework, rather than a validated coding instrument. Operational definitions, source-linked coding, and independent application remain priorities for subsequent research.

Table 4. Analytical Questions and Evidentiary Boundaries for the Exploratory 4A Framework.

Analysis Structure	Key Design Elements	Analytical Questions and Evidentiary Boundaries
Research framework	Audience participation	What does the audience do or contribute, and what data are captured? Distinguish actions, sensed signals, and model inputs.
	Art object embodiment	What is generated, and how is it presented? Separate content generation from output format, modality, and presentation effects.
	AI inquiry and response	Which AI functions and data transformations are reported? How are input, processing, and updates organized over time? Unreported implementation details and latency remain unknown.
	Aesthetic experience	What possible sensory, perceptual, and cognitive interpretations follow from the described arrangement? Distinguish artistic intention and researcher interpretation from audience evidence.
Analytical notes		Record interpretive connections, uncertainties, and proposed applications. These notes elaborate the analysis; they do not independently validate the framework.

4.3. Case Analyses

4.3.1. Unreal Pareidolia -shadows- (2023) by Scott Allen

Scott Allen's Unreal Pareidolia -shadows- is an interactive art installation that converts the shadows of everyday objects into AI-generated images and captions [59,60] (Figure 3). The documented interaction flow is as follows: A participant arranges everyday items and toys in front of a lamp on a table, where their overlapping shadows are projected onto a wall. Upon completing the configuration, the participant presses a button on the table to signal completion. The system captures the composite shadow image (a shadowgraph) via a webcam. This image is sent to an image-to-text model, BLIP-2 [48], which generates a caption displayed over the shadows on the wall in English and Japanese. Next, the caption and the shadowgraph are used as inputs (a prompt and an input image, respectively) for Stable Diffusion [47], which generates a final image. This image is projected onto the same wall as the light turns off. After a few seconds, the light turns back on, indicating that a new configuration can be made and allowing the audience to continue the interaction. Although the use of BLIP-2 and Stable Diffusion is documented, whether these pre-trained models were specifically fine-tuned for the installation is not specified in the available sources [59,60]. Table 5 presents an analysis of the work.



Figure 3. Scott Allen, Unreal Pareidolia -shadows- (2023). Image courtesy of the artist. © Scott Allen.

Table 5. 4A Framework-based Analysis of Unreal Pareidolia -shadows- (2023).

Key Design Elements	Analysis Elements	Analysis of the Work
Audience participation	Input	Everyday items and toys arranged by the audience. Shadows of the items/toys projected on a wall.
Art object embodiment	Output	Generated caption, overlaid on the wall. Final AI-generated image.
AI inquiry and response	AI technologies	Image-to-image generation using Stable Diffusion; image-to-text (caption) generation using BLIP-2. Both are available as pre-trained, open-weight models; whether either was fine-tuned for this artwork is not specified in the available documentation.
	Data processing	Projected shadow image (shadowgraph) captured by a camera and used as an input image for BLIP-2 and Stable Diffusion. Generated caption describing the shadowgraph; used as a prompt for Stable Diffusion.
	Temporality	The interaction is procedural, as the participant must press a button to confirm their input, after which captions and images are generated and displayed sequentially.

In our interpretation, the aesthetic experience of the work may center on its main concept, “pareidolia,” the psychological phenomenon of perceiving familiar patterns, such as faces or silhouettes, in random visual stimuli. While selecting and placing objects on the table is a sensory and perceptual activity, observing the resulting shadows and AI images may require cognitive effort to discover recognizable patterns in the shadowgraph and relate them to the image produced by Stable Diffusion. Notably, the AI may defamiliarize these familiar objects by producing captions or images that differ from what a participant might expect or imagine from the shadowgraph. From a technical perspective, we interpret this discrepancy in relation to an “out-of-distribution” or “out-of-domain” problem, wherein the inference inputs differ from the training domain of the model. The work can therefore be interpreted as creatively incorporating AI uncertainty and variation into its artistic structure. Although no direct audience data were collected in this study, these discrepancies may afford potential experiences of unfamiliarity, perceptual estrangement, surprise, or cognitive disruption.

Several analytical points follow. First, the work exemplifies the combined use of caption generation and image synthesis from a single input (the shadowgraph). Second, the tangible interaction positions the audience as the primary creator of the input data. Third, the explicit submission and sequential presentation separate preparation from the later AI

response. This staged organization makes the case an illustrative example of this temporal arrangement. It may afford anticipation or reflection, but those experiences were not measured in this study.

4.3.2. ReCollection (2022–2024) by Weidi Zhang and Rodger Luo

ReCollection is an interactive AI art installation that assembles synthetic collective memories from language input provided by the audience. It blurs the boundaries between remembrance and imagination through its system design and experimental visualization [24,61,62]. Figure 4 presents a representative image of the artwork. According to the artists' technical documentation [24], interaction with the AI system proceeds through four main stages. First, a participant whispers a brief personal memory into a microphone, producing approximately 10 s of recorded speech. Second, the system converts this speech to text using OpenAI's Whisper model [65]. Third, it optionally completes or structures the narrative using a GPT-4 model [66] fine-tuned on documentaries of visual memories of Alzheimer's patients and their descriptions. Finally, the resulting text prompt is fed into Stable Diffusion [47] to generate an image sequence. These images are presented as an evolving visual composition that draws on fine art techniques, such as monotype-inspired treatments and slit-scan effects. Although the AI system generally requires approximately 20 s to generate a new visualization [24], the installation maintains a continuously evolving visual display during this interval. Visual memories generated from the previous participant gradually fade and dissolve into the newly generated imagery, allowing the transition toward the current participant's "imagined memory" to unfold continuously. We interpret this design as a form of pseudo-real-time interaction: the installation incorporates that delay into a continuous visual transition between successive participants' generated memories. A summary of the analysis is presented in Table 6.



Figure 4. Weidi Zhang and Rodger Luo, ReCollection (2022–2024). Image courtesy of the artists. © Weidi Zhang and Rodger Luo.

From our interpretive perspective, the theme of the work centers on the creative reimagination of the memories of the audience—which are verbally described and often incomplete—through the lens of generative AI trained on descriptions of the memories of Alzheimer's patients. During this visual reinterpretation, the audience may engage with the aesthetic qualities of the work on multiple levels. First, whispering a personal memory may afford intimacy and self-reflection at a high level of cognition, as participants recognize they are disclosing private information to an AI, consider what to remember, and decide what to share. Observing the AI-generated image sequences may then involve perceiving the emerging visuals and interpreting how the system reimagines, extends, or

distorts that narrative by comparing it with one's own recollection. The core role of the AI can be interpreted as inferring missing details and synthesizing new visual memories from a broader corpus, using the memory of the participant as a seed for imagination. This interplay between human remembrance and machine imagination may foreground subjective interpretation and highlight how meaning emerges from the interaction between audience input and AI-driven visualization. Although no direct audience data were collected in this study, such cognitive engagement may enable potential aesthetic experiences marked by the unfamiliarity of reconstructed memories, surprise or confusion produced by distortion and reinterpretation, or empathetic resonance with the experiences of others.

Table 6. 4A Framework-based Analysis of ReCollection (2022–2024).

Key Design Elements	Analysis Elements	Analysis of the Work
Audience participation	Input	Audience's speech (personal memory) whispered into a microphone.
Art object embodiment	Output	Generative image sequences displayed as evolving visual compositions, with animation effects that may contribute to perceived immediacy; this effect is not empirically established by audience data.
AI inquiry and response	AI technologies	Speech-to-text processing using Whisper; narrative processing using a fine-tuned GPT-4-based model; text-to-image generation using Stable Diffusion. The artwork's AI pipeline runs in the cloud. Whisper and Stable Diffusion are available as pre-trained, open-weight models.
	Data processing	Text prompt generated via speech-to-text conversion and optional narrative completion.
	Temporality	Pseudo-real-time: A sequential speech-to-image pipeline includes a ~20 s generation delay, bridged by continuous transitions from previous to newly generated imagery.

Several interesting points merit discussion. First, from a technical perspective, the work employs a fine-tuned GPT-4 model trained on specific documentary data related to Alzheimer's disease. This curatorial approach to the training data aligns with the thematic focus on fragmented memories, drawing attention to both the theme and the issues surrounding the community from which the data originate. Second, Whisper's multilingual support demonstrates the capacity of AI to increase accessibility for a broader audience without substantial localization efforts [24]. Third, the work shows how voice user interfaces can be effectively applied to LLM-based AI art installations, integrating speech-to-text and text auto-completion rather than merely accepting raw prompts. Finally, the step-wise yet responsive pipeline—from speech capture to animated imagery—may support perceived continuity or immersion and afford reflection, enabling participants to consider how their narratives are transformed into synthetic, collectively inflected visual memories.

4.3.3. Visions of Destruction (2023) by Varvara Guljajeva and Mar Canet

Visions of Destruction is an interactive AI artwork that links spectatorship to environmental changes within synthetic landscapes [18,43,63]. The documented interaction flow is as follows. The installation continuously renders landscape scenes using Stable Diffusion [47]. As shown in Figure 5, when a viewer looks at specific regions on the display, an eye-tracking system tracks their gaze and aggregates fixations into a mask. The system then applies AI inpainting to those masked areas, replacing healthy terrain with

polluted, degraded motifs selected from a topic-relevant prompt set. If no gaze is detected for a specific period, the scene gradually regenerates through ongoing text-to-image synthesis and smooth latent interpolations. The simple act of looking becomes a catalyst for damage, while looking away enables recovery. The system uses a pre-trained Stable Diffusion model running locally on a machine with an NVIDIA RTX 4090 GPU [18]. Table 7 summarizes the described arrangement.



Figure 5. Varvara Guljajeva and Mar Canet, Visions of Destruction (2023). Image courtesy of the artists. © Varvara Guljajeva and Mar Canet.

Table 7. 4A Framework-based Analysis of Visions of Destruction (2023).

Key Design Elements	Analysis Elements	Analysis of the Work
Audience participation	Input	Viewer’s gaze and head position, captured via an eye tracker.
Art object embodiment	Output	Generative landscapes, either continuously updated with smooth animation or locally modified via AI inpainting.
AI inquiry and response	AI technologies	Text-to-image generation using Stable Diffusion (with preset landscape prompt sets); latent-space interpolation used to create smooth animations; image inpainting using Stable Diffusion, guided by gaze-derived masks and predefined prompts. Stable Diffusion is available as a pre-trained, open-weight model. The artwork runs it locally on a machine with an NVIDIA RTX 4090 GPU.
	Data processing	Aggregated gaze mask, computed from gaze fixations with a slight delay.
	Temporality	Continuous gaze sensing is combined with gaze aggregation, inpainting updates, and ongoing landscape presentation. Real-time sensing does not establish immediate inference; The available documentation does not report measured end-to-end latency.

In our interpretation, the aesthetic experience of the work may emerge from a tightly coupled loop between seeing and altering. On a sensory level, viewers encounter luminous, slowly shifting panoramas punctuated by local, gaze-triggered transformations. Perceptually, participants may recognize a contingent rule: wherever their attention lingers, decay spreads. Cognitively, the piece may invite reflection on the observer effect and human responsibility for ecological harm. The act of averting one’s gaze to allow recovery may produce a deliberate tension between curiosity and restraint, turning self-regulation into an ethical and conceptual gesture. Through its generative AI-driven cycle of destruc-

tion and recovery, the artwork may afford potential experiences of shock and fear, as well as relief through restoration. In repeatedly exposing audiences to the consequences of their own participation, the work may invite reflection on the ethical implications of their actions and may foster experiences of ethical awareness, empathy, and self-reflection.

Several aspects merit attention. First, the installation demonstrates a coherent integration of gaze tracking with diffusion-based image generation and inpainting. Latent interpolation sustains continuity, while inpainting localizes change without resetting the entire scene. Second, the interaction model can also be interpreted as minimal, with gaze serving as both a sensor and interaction mechanism, potentially lowering the barrier to entry while maintaining participant agency. Third, the temporal design balances immediacy and delay. Specifically, real-time tracking provides responsiveness, whereas the slight lag in mask aggregation and inpainting may encourage anticipation and reflection rather than frantic scanning. Finally, the regeneration mechanism may reframe noninteraction as a meaningful choice. Withholding attention can be interpreted as an act of care, which may deepen the conceptual reading of the work and broaden the range of audience experiences within the system.

5. Discussion

This discussion brings together three design-oriented considerations developed through the exploratory reviews and subsequent case analyses: AI temporality, modality conversion, and AI-mediated aesthetic experience. In relation to the 4A framework, AI temporality concerns the timing and sequencing of audience participation, AI processing, and output presentation; modality conversion concerns the transformation of audience input through computational processing into artistic output; and AI-mediated aesthetic experience concerns the potential sensory, perceptual, and cognitive experiences afforded by these arrangements. The temporal contrasts informed case selection and are further elaborated here, while modality conversion and AI-mediated aesthetic experience were developed as additional strands through subsequent analysis and writing. Together, these considerations inform exploratory design propositions rather than empirically validated design principles.

The case analyses also raised ethical and practical considerations that extend beyond these three design-oriented aspects. Accordingly, the discussion addresses potential ethical issues associated with AI-mediated audience interaction and practical and technical constraints that may affect the deployment and sustained operation of AI-driven installations in exhibition environments. As these concerns were not a primary focus of the framework or the case analyses, their treatment is selective and exploratory rather than comprehensive, focusing on issues that arose from examining the selected works.

5.1. AI Temporality in System Design and User Interaction

In interactive installations incorporating AI, the time from user input to system output is critically dependent on the temporality of the AI technologies involved. This temporality becomes a key determinant of both system design and user interaction, ultimately shaping the aesthetic experience. System design must account for inference delays, which vary with factors such as model architecture, deployment mode, and available computing resources. These constraints raise several questions: How should an installation handle delay or latency introduced by AI inference processes? How might artists leverage or mitigate the temporal constraints of AI models to achieve aesthetic goals? How do the temporal characteristics of AI systems influence artistic practice and the experience of the viewer? The following three approaches elaborate the provisional temporal contrasts that informed case selection. They describe arrangements of interaction rather than measured

speed classes, can overlap within a single work, and do not independently validate a three-part temporal taxonomy.

Approach 1, Procedural, Event-driven, or Delayed Interaction: A system can intentionally structure multiple steps between user input and AI output. This approach may be effective when real-time inference is infeasible on local hardware or when models run remotely via application programming interfaces (APIs) or other networked services. Interaction can be procedural, with step-by-step instructions guiding input preparation; event-driven, with graphical user interface elements enabling users to enter, modify, and confirm input; or delayed, with the system visualizing processing while computation proceeds. For example, in *Unreal Pareidolia -shadows-*, participants first arrange physical objects and then press a button to confirm their arrangement [59,60]. A caption for the shadowgraph must be generated by BLIP-2 [48] before Stable Diffusion [47] can synthesize the image, which introduces a delay. The system displays the caption before the generated image [59,60]. Drawing on the distinction between subjective and objective temporality [67], this sequencing may reduce “felt time” by cueing expectations about the forthcoming output, without necessarily shortening the “measurable time” required for image generation.

Figure 6 presents a schematic design proposition informed by this temporal arrangement. Progress indicators, countdowns, or preview states are possible extensions for making computational stages legible. Where submissions are queued or batched, indicating queue position may help manage expectations and reduce frustration. More broadly, staging and interim feedback may transform AI-related latency from a disruption into a temporal structure that supports anticipation and reflection.

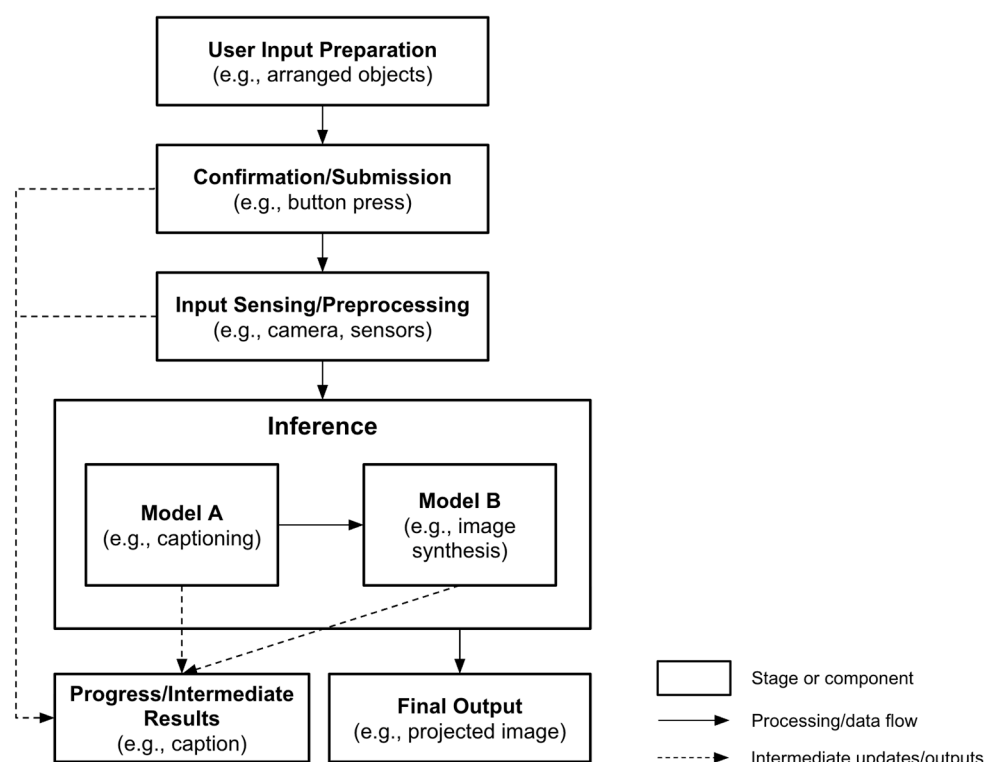


Figure 6. Procedural, event-driven, or delayed interaction. A schematic design proposition organizing input preparation, explicit submission, sequential AI processing, and output presentation into distinct stages. Intermediate results or progress feedback could make computational stages legible while the final output is being generated.

Approach 2, Pseudo-Real-Time Interaction: As shown in Figure 7, AI-based installations may maintain a sense of temporal continuity even when generation itself is not immediate. In ReCollection, participants provide approximately 10 s of voice input, after which the cloud-based AI system generally requires around 20 s to generate a new visualization [24]. Rather than leaving the display static during this interval, the installation gradually fades and dissolves visual memories generated from the previous participant's input into imagery generated from the current participant's input [24]. The installation therefore does not eliminate computational latency; instead, it incorporates the waiting period into an evolving visual transition. This approach suggests that computational delay can be treated as part of an artwork's temporal composition rather than solely as a technical limitation. Continuous or transitional visual states may bridge separate generative responses and maintain continuity across successive interactions.

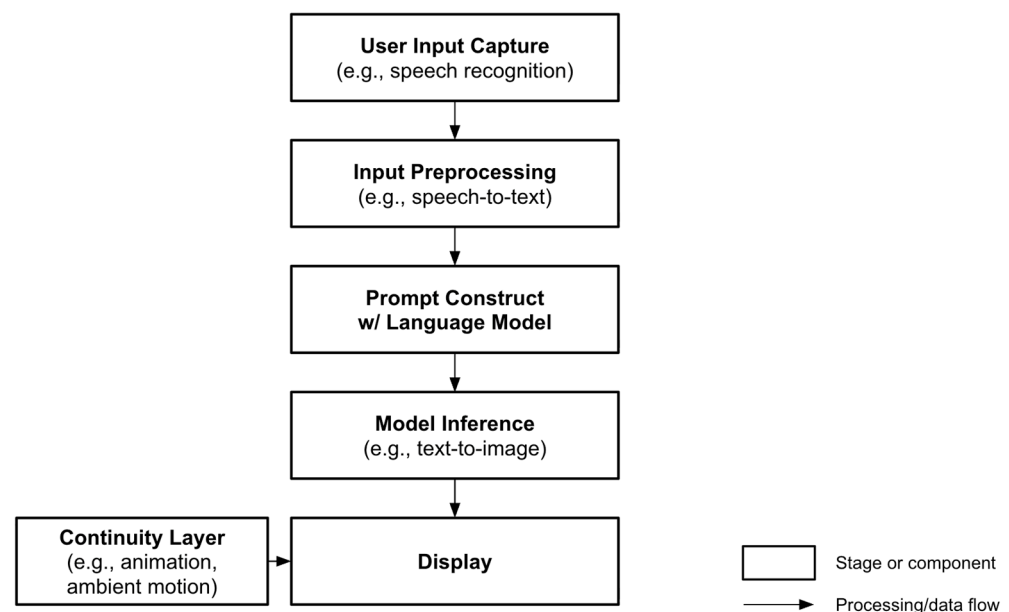


Figure 7. Pseudo-real-time interaction. A schematic design proposition combining a sequential speech-to-image pipeline with a continuity layer, such as animation or ambient motion, to bridge computational delay. “Pseudo-real-time” describes continuous presentation during delayed processing, rather than instantaneous inference or an empirically established perception of immediacy.

Possible extensions include streaming partial results, progressive refinement, prefetching predictable assets, and double-buffering visual states for smooth transitions. Continuous ambient soundscapes, low-salience auditory layers, or background motion may also be used while new results are produced. These strategies may contribute to a sense of liveness and immersion by keeping the installation perceptually active during computational delay, while potentially shifting participants' attention toward the evolving presentation of incoming AI-generated content.

Approach 3, Continuous or Real-Time Interaction: Installations can employ AI capabilities that support near-real-time output while accommodating continuous user input. In this approach, user actions directly generate new content or modify the existing content in place. As illustrated in Figure 8, AI can function as a tool that produces a steady stream of frames for smooth animation or an agent that alters content using algorithms such as inpainting. In *Visions of Destruction*, real-time gaze tracking drives Stable Diffusion [47] inpainting within specific regions, while latent-space interpolation sustains smooth landscape animation [18]. Together, these mechanisms maintain real-time interactivity.

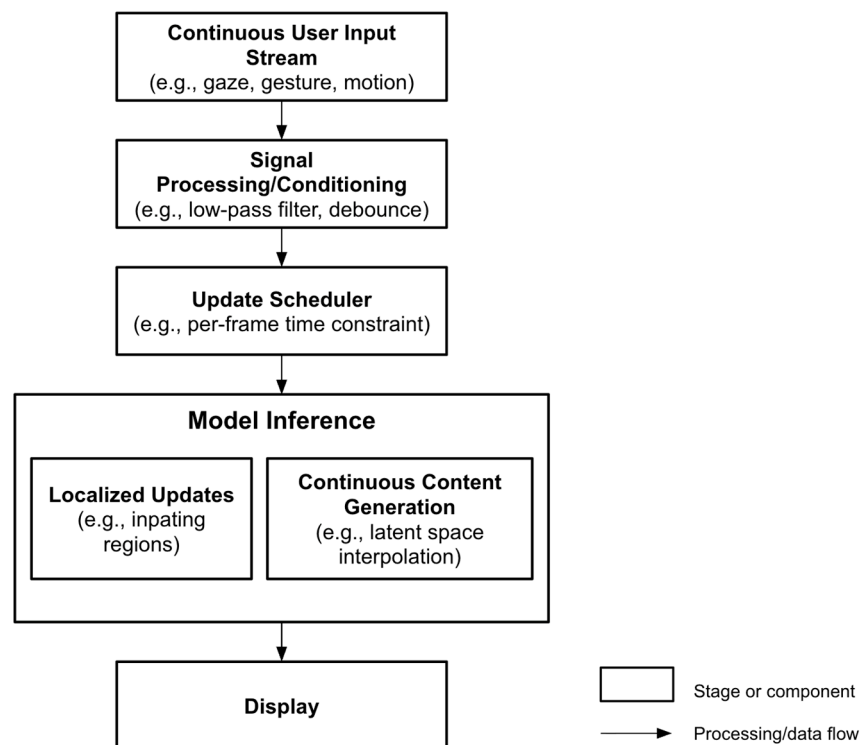


Figure 8. Continuous or real-time interaction. A schematic design proposition coupling continuous participant input with signal conditioning, update scheduling, and localized modification or ongoing content generation. Sensing, AI processing, and display updates may operate at different rates within this arrangement.

Potential design extensions include throttling update rates, applying temporal easing, restricting updates to local regions within a frame-time budget, and using low-pass filtering or debounce constraints on sensor input. Caching prompts or seeds and considering distilled or on-device models are further implementation possibilities. This immediate interaction may support a sense of agency, contingency, and presence. The tight input–output coupling could make the relationship between participant actions and system responses more legible, support improvisation, and allow feedback to function as an expressive element of the interaction.

Taken together, these cases suggest that AI temporality can function as both a technical constraint and an expressive material shaping interaction and meaning. Procedural or delayed sequencing may support anticipation; pseudo-real-time orchestration may maintain continuity across computational delays; and immediate input–output coupling may strengthen the legibility of participant agency and presence. More broadly, sequencing, continuity cues, and stabilized updates can incorporate computational latency into the temporal structure of an artwork, potentially contributing to its aesthetic character.

5.2. Modality Conversion in AI Interaction

Across the three case studies, interaction with generative AI extends far beyond simple text entry. Generative AI pipelines often rely on joint text–image representations and transformer-based models to achieve multimodal generation (e.g., text-to-image and text-to-sound). However, text need not be the primary input modality of the audience. Instead, systems can convert tangible, verbal, or embodied signals into internal representations that condition or contribute to the generation process. This conversion chain may be central to how human and machine agency, artwork meaning, and interaction timing are experienced.

Approach 1, Tangible Interaction: Tangibility can be an integral mode of interaction in AI-based installations, coupling material manipulation with computational imagination. In *Unreal Pareidolia -shadows-*, participants arrange physical objects to produce a shadowgraph, which is then captured by a camera, captioned by an image-to-text model, and used to condition image synthesis [59,60]. This approach transduces a tactile action into a visual latent representation and a textual prompt, which together steer the output. Figure 9 illustrates this conversion from tangible input to the desired output modality.

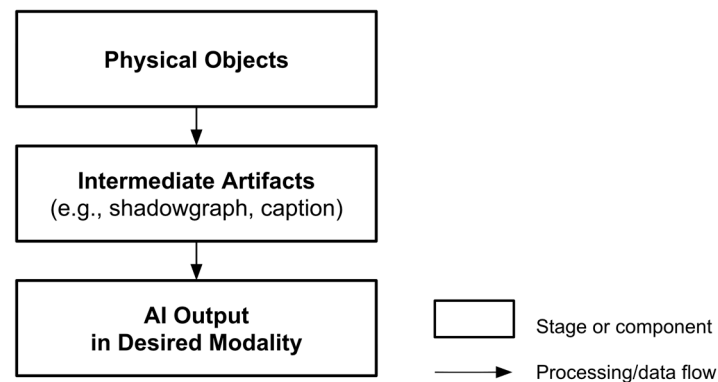


Figure 9. Modality conversion through tangible interaction. A schematic pathway from physical object arrangements to intermediate artifacts, such as shadowgraphs and captions, that condition AI-generated output. The diagram illustrates how material manipulation can be translated into representations used by generative models.

Tangible interaction could support modality conversion beyond image synthesis. For instance, drawings depicting the mood of a participant could serve as inputs for image-to-sound or image-to-music generation. Physical collages could be segmented into layout prompts for story or caption generation. Clay or modular sculptures scanned in situ could be used to initialize 3D scene or object synthesis via image-to-3D pipelines. Vectorized textile or weaving artifacts could be used to condition the generation of materials and textures for 3D assets. Additionally, tangible user interfaces with buttons, sliders, or knobs could enable participants to engage directly with AI content generation by manipulating controllable parameters rather than serving only as providers of raw input, supporting iterative, hands-on exploration and fine-grained manipulation of the output.

Approach 2, Verbal Interaction: As illustrated in Figure 10, verbal interfaces can position speech as the primary conduit into generative AI pipelines by converting acoustic signals into textual or semantic representations. In *ReCollection*, the system records a whispered memory, performs speech recognition, optionally restructures the narrative with an LLM, and then drives image generation [24]. Thus, speech becomes a narrative text that conditions a visual output, demonstrating how conversational AI interfaces can be adapted to installation contexts.

Verbal interaction has broader applications. For example, spoken scene descriptions could be translated into structured scripts for text-to-video generation; prosodic features (e.g., tempo and intensity) could modulate style parameters for music or affective color palettes for images; and multilingual speech could be normalized into prompts that control 3D scene synthesis or texture generation. Moreover, voice-driven refinement loops, wherein the system exposes interim transcripts or candidate paraphrases for confirmation, could support iterative engagement without restarting the process to provide new input.

Approach 3, Embodied Interaction: As shown in Figure 11, embodied interfaces can mobilize sensorimotor signals as continuous control fields that directly modulate generative AI processes. In *Visions of Destruction*, gaze is tracked in real time, aggregated

into spatial masks, and used to localize inpainting within a continuously generated landscape [18]. Thus, embodied signals function as control layers over the evolving scene.

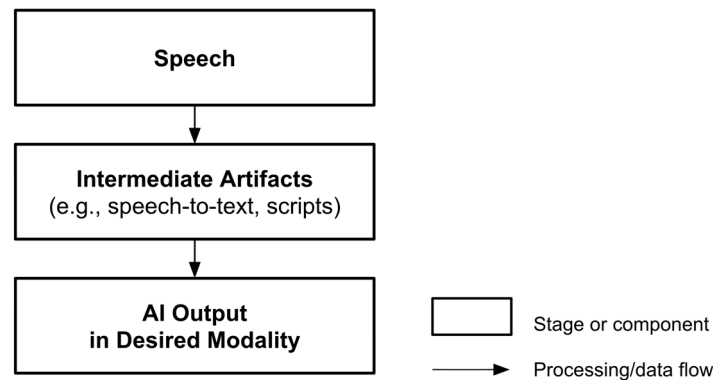


Figure 10. Modality conversion through verbal interaction. A schematic pathway in which speech is converted into intermediate textual representations, such as transcripts or structured narratives, that condition AI-generated content in a chosen output modality.

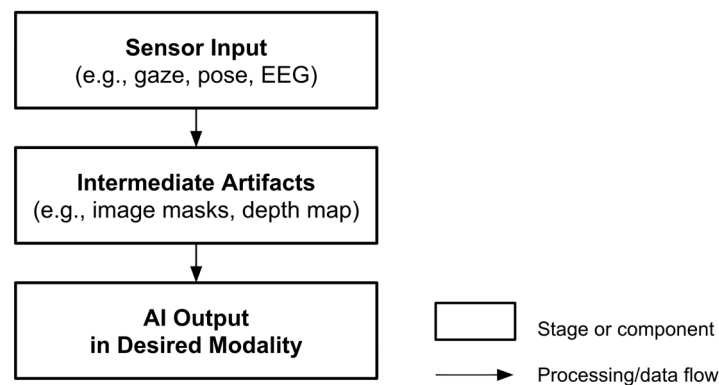


Figure 11. Modality conversion through embodied interaction. A schematic pathway from sensor inputs, such as gaze, body pose, or electroencephalography (EEG) signals, through intermediate representations to AI-generated output. The mappings illustrate possibilities for translating bodily input into representations that could guide content generation or modification.

Beyond gaze, designers could incorporate various embodied interactions. For instance, whole-body poses could guide character synthesis or choreography in video generation; proxemic distance and orientation could adjust guidance strength, field of view, or depth of field in image synthesis; and motion-captured props could instantiate or reposition 3D objects within text-conditioned scene construction. Physiological inputs, including heart rate, respiration, electroencephalogram (EEG) signals, or electromyography signals, could modulate tempo, contrast, saturation, or noise schedules in outputs that track body state. To sustain expressive control, systems could expose calibrated mappings between embodied signals and generative parameters, allowing participants to continuously shape content rather than merely trigger it.

These modality conversions suggest several considerations for interaction design. First, the accuracy, lossiness, and bias of each converter may influence how faithfully intent is conveyed. Second, the granularity of control depends on the representation; continuous masks may afford different forms of agency from those afforded by discrete prompts. Third, transparency regarding intermediate artifacts (e.g., captions, transcripts, or masks) could make conversion a collaborative step, allowing participants to steer the system before final rendering. Finally, because conversion stages impose their own latencies, or-

chestration strategies such as progressive refinement, preview states, or continuous background motion could help manage perceived delay.

By foregrounding modality conversion, interactive AI artworks could move beyond prompt-centric interfaces toward plural, situated forms of engagement that align audience capabilities with the representational affordances of generative models. Multimodal redundancy may support accessibility by providing alternate pathways for participation, whereas mixed-initiative repair loops may mitigate conversion errors.

5.3. AI-Mediated Aesthetics

Building upon the proposed analytical framework, this study considers how AI may mediate potential forms of aesthetic experience in interactive art. Across the three illustrative cases, three design-oriented approaches are considered: (1) exploring differences between human and machine perception, (2) enabling personalized aesthetic experiences, and (3) sustaining or expanding aesthetic interaction. These approaches are interpretive considerations rather than empirically confirmed audience effects, and the aesthetic responses discussed below should therefore be understood as anticipated or potential experiential affordances.

Approach 1, Exploring Differences between Human and Machine Perception:

Machine perception processes sensory data through computational representations and probabilistic inference, differing from human perception grounded in embodied experience, cultural context, and accumulated knowledge [68–72]. Because AI outputs are constrained by training data and computational processes, differences between human and machine perception may create unexpected or unfamiliar representations. In *Unreal Pareidolia -shadows-*, for example, AI transforms silhouette information into textual descriptions and generated images that may diverge from participants' expectations [59,60]. Such discrepancies can be interpreted as a form of defamiliarization and may afford potential experiences of unfamiliarity, surprise, or perceptual ambiguity. In this way, differences between human and machine perception may serve as aesthetic resources in AI-mediated interactive art. Although machine vision is not inherently aesthetic, its integration may therefore provide artists with opportunities to explore tensions between human and non-human cognition and expand creative possibilities through algorithmic mediation.

Approach 2, Enabling Personalized Aesthetic Experiences: AI may support personalized aesthetic experiences by responding dynamically to user-specific data. Unlike traditional interactive installations that often rely on predefined scenarios, AI-based systems can process individual inputs and generate context-sensitive outputs that may be personally or emotionally relevant. In *ReCollection*, multilingual AI models process audience voices and transform personal memories into individualized visual and auditory compositions [24]. These AI-mediated representations may afford intimate or reflective engagement, although such experiential effects were not measured through audience data. More broadly, culturally diverse datasets may provide opportunities to encounter perspectives beyond one's own and potentially support cross-cultural empathy. These possibilities illustrate how AI may support participant-specific and context-aware forms of interactive art.

Approach 3, Sustaining or Expanding Aesthetic Interaction: AI-driven interactivity may support continuous and evolving engagement, rather than discrete, one-time interactions. Traditional interactive systems often rely on fixed content, whereas AI allows artworks to adapt dynamically over time, potentially encouraging sustained participation and iterative exploration. *Visions of Destruction* illustrates how continuously changing AI-mediated responses may encourage repeated interaction between audience attention and the evolving artwork [18]. Conversely, in *Unreal Pareidolia -shadows-* [59,60], AI un-

predictability may produce responses that extend beyond both the artist's intention and audience expectations. Such responses can be interpreted in relation to prior notions of unintended experience [10,16], as seen in earlier works such as Legible City [73] and Rain Room [74], wherein participants deliberately subverted system constraints. From a design perspective, AI may therefore provide a dynamic space between artistic control and emergent interaction, suggesting the need to consider how system autonomy and user agency are balanced within evolving aesthetic experiences.

Across these three approaches, AI can be understood as a co-creative agent that may extend perceptual possibilities, support personalized interaction, and sustain evolving forms of engagement. The three illustrative cases suggest potential conditions for emergent, affective, and reflective experiences rather than empirically confirmed audience effects. These interpretations provide an exploratory basis for considering aesthetic interaction as an important dimension of human–AI collaboration in interactive art.

5.4. Ethical Considerations

AI-driven interactive art may raise ethical concerns when audience participation involves the collection, interpretation, or transformation of human-generated or personal data. Key considerations include privacy, informed consent, algorithmic bias, data ownership, authorship, agency, and the appropriate representation of sensitive information [25,26]. These concerns may become particularly relevant when AI systems interpret audience input, generate representations based on personal data, or collect behavioral and physiological information.

Several of these ethical concerns can be considered in relation to the selected cases. Algorithmic bias and the limitations of AI interpretation are relevant when AI-generated outputs represent or transform audience input; for example, in Unreal Pareidolia -shadows- [59,60], generated captions and images may differ from the participant's original input or intended interpretation. Privacy, data ownership, and the representation of sensitive information become important when personal or sensitive data are incorporated into artistic production, as illustrated by ReCollection and its use of memory-related material associated with individuals with Alzheimer's disease [24]. Informed consent and audience-data collection are particularly relevant to sensor-based interaction; Visions of Destruction, for example, uses gaze tracking as part of its interaction mechanism, illustrating the need to consider how behavioral data are collected and used [18]. These examples suggest that ethical concerns in AI-mediated art may vary according to the types of data, AI processes, and interaction mechanisms involved.

The present study does not empirically or systematically evaluate the ethical impacts of these artworks; rather, these issues are identified as potential considerations arising from the case analyses. Ethical considerations should therefore be taken into account when designing and deploying AI-driven interactive installations, particularly where personal data, behavioral input, or algorithmically generated representations are involved. Further research is needed to examine these issues systematically and to develop appropriate ethical guidance for AI-driven interactive art.

5.5. Practical and Technical Constraints

AI-driven interactive installations may face practical and technical constraints involving system reliability, processing latency, model unpredictability, sensing accuracy, accessibility, and hardware or network dependencies. The documented configurations of the three cases suggest distinct potential challenges. In Unreal Pareidolia -shadows-, image capture, caption generation, image synthesis, and projection form a sequential pipeline in which interruptions at one stage could disrupt subsequent stages [59,60]. ReCollection

requires coordination between its cloud-based speech-to-image pipeline and ongoing visualization. The artists report a generation time of approximately 20 s, which varies with internet connection speed [24]. For *Visions of Destruction*, the coupling of gaze sensing with generative updates raises practical considerations concerning processing coordination and tracking consistency across participant positions, movements, and environmental conditions [18]. These are potential constraints inferred from the system's configuration, rather than observed failures.

Across and beyond these cases, excessive latency, component failures, and unavailable network services may interrupt interaction, while unpredictable model outputs may challenge artistic consistency. However, unpredictability should not automatically be equated with technical failure, as variation may also serve as an aesthetic resource.

Exhibition planning should therefore consider adequate computational and network resources, sensor calibration, lighting and sound conditions, interface accessibility, and procedures for monitoring, maintenance, and recovery. Alternative input options or staff-supported participation could accommodate audiences who cannot readily use the required interaction modes. These considerations are selective and exploratory, as technical performance and accessibility were not primary concerns of the framework or analyses. Further research should evaluate how these constraints affect installation operation and whether the proposed measures support interaction continuity and accessibility in actual exhibition environments.

6. Conclusions

6.1. Implications

This study proposed the 4A framework as an exploratory analytical framework for examining how AI technologies shape interaction processes and aesthetic experience in interactive artworks. The framework comprises three key entities (audience, art object, and AI) and four design elements (participation, embodiment, AI inquiry and response, and aesthetic experience), extending existing approaches to account more explicitly for the mediating role of AI. The analysis of three illustrative cases—*Unreal Pareidolia -shadows-*, *ReCollection*, and *Visions of Destruction*—suggests that AI may function not merely as a technical tool but also as a critical mediating element that connects artistic intention, system architecture, and audience participation. In relation to the 4A framework elements, the selected cases provide a basis for considering AI temporality, modality conversion, and AI-mediated aesthetic experience as design propositions rather than as generalizable findings. Based on these observations, the following implications are proposed as exploratory considerations for the analysis and design of AI-driven interactive art.

Creative Use of Cognitive Gaps: Discrepancies between human and AI reasoning may be considered potential aesthetic resources rather than mere limitations. They may evoke ambiguity, surprise, and interpretive openness, thereby creating opportunities for audience imagination and engagement. In AI-driven interactive installations, these cognitive contrasts may support aesthetic possibilities that operate at a perceptual or conceptual level distinct from human cognition. At the same time, these gaps may also function as points of failure, potentially leading to confusion or frustration rather than aesthetic engagement. When carefully mediated through appropriate interaction design strategies, such as procedural or delayed interaction, they may instead contribute to novel forms of aesthetic experience.

Integrated Design of Multimodal Inputs: Combining multimodal and embodied input methods, such as voice, gaze, and bodily movement, may broaden participation and support immersive engagement. These multimodal interfaces offer different physical and sensory ways of contributing to an artwork. Their integration requires attention to how

heterogeneous inputs are interpreted by AI systems and mapped to artistic outputs. Complementary inputs could provide distinct forms of expressive control, while redundant pathways could accommodate varied participant capabilities and help sustain interaction when one input channel is unavailable or unreliable. Such arrangements may support diverse forms of artistic expression, with their experiential effects depending on the installation's design and context.

Embracing Implicit and Nonverbal User Behaviors: Audience input should not be limited to explicit actions. Subtle, unconscious, or nonverbal gestures may also provide meaningful input for AI-mediated artistic responses. Such approaches may create opportunities for serendipity and improvisation and expand the range of aesthetic possibilities emerging from interactions between audiences and AI-driven installations. However, their effectiveness and experiential implications require further examination across different interaction contexts and audience populations.

Maintaining Artistic Consistency Amid Uncertainty: AI systems could be designed to support artistic intent and maintain coherent behavior when audience input is excessive, insufficient, or ambiguous. Such design may help sustain aesthetic continuity under uncertain interaction conditions. At the same time, audience interpretations and responses may extend beyond the artist's intentions. Artists could therefore consider how to maintain a recognizable creative direction while incorporating selected forms of unpredictability into the work's design, allowing unexpected interactions to contribute to its aesthetic possibilities.

Together, these insights suggest design considerations and creative approaches for AI-based interactive art installations. They also point to possibilities for exploring how AI-mediated interaction may shape aesthetic experience within specific artistic and interactional contexts.

6.2. Limitations and Future Work

The limited number of analyzed artworks represents an important limitation of this study. The three selected works were used as illustrative cases to examine the application of the 4A framework and should not be considered representative of AI-driven interactive art as a whole. Accordingly, the case analyses are intended to support exploratory examination rather than broad empirical generalization. Future research should apply the framework to a wider and more diverse range of AI-based interactive artworks, exhibition contexts, and interaction configurations to examine its transferability, refine its analytical dimensions, and assess its potential applicability across different contexts.

Another limitation lies in the treatment of aesthetic experience. Because it was not feasible within the scope of this research to collect primary empirical data, such as audience interviews, surveys, or behavioral measurements, the analysis relied on the official descriptions of the artworks, related publications, and publicly available video documentation of the installations. Accordingly, interpretations of aesthetic experience should be understood as potential or analytically inferred experiential affordances rather than as empirically validated accounts of audience experience. Future research should incorporate audience-centered empirical studies to examine how these experiential possibilities are perceived by participants and to evaluate the framework through direct audience evidence.

The technical analysis of the selected artworks was also constrained by the availability of public documentation. Although the detailed analysis provides additional information on the AI models, data processing, and implementation characteristics where documented, detailed information regarding model architectures, datasets, training procedures, and implementation was not consistently available across all cases. The study therefore limited its

technical descriptions to verifiable information reported in publicly accessible sources and explicitly distinguished documented information from the authors' analytical interpretations. Future research could extend this analysis through access to more detailed technical documentation or direct collaboration with artists and developers.

Finally, ethical considerations remain important in AI-driven interactive art, particularly when artworks involve personal or sensitive audience data and AI-mediated artistic production. Although the present study discusses relevant ethical issues, it does not systematically or empirically evaluate their effects across different artistic contexts. Future research should therefore further examine how ethical considerations can be incorporated into the analysis and design of AI-driven interactive art while preserving its experimental and aesthetic potential.

Author Contributions: S.P.: Conceptualization, Funding acquisition, Resources, Methodology, Visualization, Writing—original draft, Writing—review and editing; J.-h.O.: Conceptualization, Project administration, Visualization, Writing—original draft, Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was undertaken thanks in part to funding from the Connected Minds Program, supported by Canada First Research Excellence Fund, Grant #CFREF-2022-00010.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding authors.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT (OpenAI; models ranging from o1 to GPT-6 Astra) for methodological discussion, literature-search assistance, drafting and revising selected passages, language editing, and proofreading. The authors have reviewed and edited the generated output as necessary and take full responsibility for the content of this article. This work was supported by BK21 FOUR (Fostering Outstanding Universities for Research) (No. 5199990914048).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial intelligence
ML	Machine learning
VAE	Variational autoencoder
LLM	Large language model
DL	Deep learning
IEEE	Institute of Electrical and Electronics Engineers
CVF	Computer Vision Foundation
CVPR	Computer vision and pattern recognition
ACM	Association for Computing Machinery
HCI	Human–computer interaction

Appendix A

Table A1 lists the 13 artworks identified as interactive AI installations in the CVPR 2024 AI Art Gallery.

Table A1. AI-based Interactive Art Installations in the CVPR 2024 AI Art Gallery.

No.	Title	Artist(s)	Year
1	Unreal Pareidolia -shadows-	Scott Allen	2023
2	Visions of Destruction	Varvara & Mar	2023

Table A1. *Cont.*

No.	Title	Artist(s)	Year
3	AI Nūshu (AI 女书)	Yuqian Sun	2023
4	The Endless Collaborative Book: Wushu	Aven Le Zhou	2023
5	Connections	Yamin Xu	2023
6	Neuracappella	Zihou Ng	2023
7	Speculative Evolution, Prototype 1	Marc Lee	2024
8	YouTube Mirror	Sihwa Park	2022–2023
9	Gender Tapestry	J. Rosenbaum	2023–2024
10	ReCollection	Weidi Zhang, Rodger Luo	2022–2024
11	On The Existence of Self-Identity	Julia Chylak, Mateusz Błajda, Eryk Imos, Beata Bajno, Paulina Wachnicka	2023
12	Beyond Characters: The Unseen Labyrinth	Mingyong Cheng, Zetao Yu	2023–2024
13	LVX-1	LB ²	2024

Appendix B

This appendix summarizes the six exploratory artwork reviews that informed framework development. The summaries retain the original categories of input, output, interaction temporality, AI technologies, and key points, including the technical, interaction-related, and aesthetic observations recorded in the notes. They document preliminary descriptions and interpretations rather than complete applications of the final 4A framework. Temporal labels reflect initial qualitative assessments rather than independently measured latency, and aesthetic observations are interpretations rather than established audience responses. Uncertainties recorded during the reviews are retained.

Appendix B.1. Unreal Pareidolia -shadows- (2023)

Input: Everyday items and toys arranged by participants, the resulting shadows, a camera-captured shadowgraph, and an AI-generated caption used as a prompt for image generation.

Output: A caption describing the shadowgraph and an AI-generated image superimposed on the shadows through projection mapping.

Interaction temporality: The initial review described the interaction as “non-real-time” because generation followed a button press. The artwork description quoted in the notes, however, described image and caption generation as occurring in real time. These descriptions were recorded without quantitative latency measurements.

AI technologies: Stable Diffusion for image-to-image generation and BLIP-2 for image captioning.

Key points: The review highlighted the combined use of caption and image generation from the same shadowgraph and the audience’s tangible contribution to creating input data. Aesthetic observations concerned the defamiliarization of familiar objects and the use of unexpected captions and images in relation to pareidolia. The interpretation that these discrepancies involved uncertainty or an “out-of-domain” problem was recorded as an exploratory analytical explanation.

Appendix B.2. ReCollection (2022–2024)

Input: Participants' speech captured through a microphone and converted into textual prompts.

Output: Generative image sequences modified through visualization techniques drawing on monotype printmaking and slit-scan photography.

Interaction temporality: Initially characterized as "near immediate," with speech processing followed by image generation and display. The review interpreted real-time animation as contributing to an appearance of immediacy, while leaving the role of a button press in initiating voice recording uncertain.

AI technologies: Whisper for speech-to-text processing, GPT-4 for text auto-completion, and Stable Diffusion for text-to-image generation. The notes referred to a fine-tuned transformer-based model, but its identification as a fine-tuned GPT-4 variant remained tentative.

Key points: The review emphasized the combination of a voice interface, multilingual speech processing, and participant-specific narrative input. It also noted the use of documentary material concerning the visual memories of people with Alzheimer's disease. The aesthetic interpretation focused on the relationship between fragmented personal recollections and AI-generated visual imagination, through which details could be supplemented or transformed into synthetic memories.

Appendix B.3. Beyond Characters: The Unseen Labyrinth (2023–2024)

Input: Participants' gaze.

Output: AI-generated fairy-tale text, associated images, and soundscapes. The review explicitly questioned whether the accompanying images were AI-generated.

Interaction temporality: The notes described real-time webcam tracking of gaze and head movement, gaze-responsive distortion of text, and continuous fairy-tale generation. Whether gaze also affected the images and sound remained unclear.

AI technologies: Fairy-tale text generation and a provisionally identified text-to-image function; specific models were not recorded, and the image-generation process remained uncertain.

Key points: The review focused on the artistic interpretation of reading challenges associated with dyslexia and described AI as a tool operating as a separate module within the installation. The quoted artwork description situated the transformations of text, imagery, and sound within an intention to communicate dyslexic perception and foster empathy; these were artistic aims rather than audience effects established by the review.

Appendix B.4. Visions of Destruction (2023)

Input: Participants' gaze.

Output: Generative landscape images that are continuously updated or partially modified through smooth animation.

Interaction temporality: The review distinguished continuous landscape generation when no viewer was detected, real-time gaze sensing, a small delay in aggregating gaze into a mask, and subsequent inpainting of the masked region using artist-defined prompts. AI technologies: Stable Diffusion for text-to-image generation and inpainting, together with latent-space navigation and interpolation between prompts to produce smooth transitions.

Key points: Technical observations emphasized animation through interpolation and the use of gaze to guide localized image modification. The aesthetic interpretation connected gaze-induced landscape destruction and regeneration during viewer absence to the

artists' metaphor of human environmental impact. The notes treated looking and looking away as meaningful components of the artwork's conceptual arrangement.

Appendix B.5. AI Nüshu (AI (女书) (2023))

Input: An image of participants and their surroundings, an AI-generated caption, and Nüshu scripts.

Output: The captured image and caption, a generated poem, encoded characters in the AI Nüshu dictionary, newly generated symbols, and guessed character meanings.

Interaction temporality: The review described single-image capture rather than continuous input. Subsequent processes were characterized as almost real-time, with some delay before the resulting symbol was visualized.

AI technologies: BLIP for image captioning, GPT-4 for poem generation, BERT-Chinese-Base for language embeddings, and principal component analysis for dimensionality reduction.

Key points: The review emphasized the visualization of a speaker-encoding and listener-decoding process as a language-learning game. It described direct audience interaction as limited, while identifying the transformation of audience imagery into new Nüshu-inspired symbols as an aesthetic point of interest. The quoted artwork description situated this language-generation process within a non-Western feminist context.

Appendix B.6. Wushu (2023)

Input: Participants' body and hand movements.

Output: Captured images of participants, calligraphic ink strokes and effects, and generated bilingual English–Chinese characters arranged in a book-page layout.

Interaction temporality: Real-time hand tracking and calligraphic effects were distinguished from delayed text generation.

AI technologies: The recorded artwork description identified two customized neural networks and a "Motion Caption" process. The review tentatively interpreted this as image-to-text generation or image captioning, while noting that the implementation was not clearly described.

Key points: The review identified separate "writing" and "reading" stages: participants first produce motion-based calligraphic strokes and subsequently examine generated text. It highlighted the transformation of bodily movement into calligraphy and textual content. A comparison with Unreal Pareidolia -shadows- raised the possibility of repurposing model uncertainty or input-domain mismatch as an aesthetic resource, but this remained tentative because the training process and dataset were unclear.

References

1. Kluszczyński, R.W. Strategies of interactive art. *J. Aesthet. Cult.* **2010**, *2*, 5525. [\[CrossRef\]](#)
2. Kwastek, K. *Aesthetics of Interaction in Digital Art*; The MIT Press: Cambridge, MA, USA, 2013. [\[CrossRef\]](#)
3. Lopes, D.M.M. The ontology of interactive art. *J. Aesthet. Educ.* **2001**, *35*, 65–81. [\[CrossRef\]](#)
4. Bruder, K.A.; Uçok, O. Interactive art interpretation: How viewers make sense of paintings in conversation. *Symb. Interact.* **2000**, *23*, 337–358. [\[CrossRef\]](#)
5. Ahmed, S.U. Interaction and interactivity: In the context of digital interactive art installation. In *Human–Computer Interaction. Interaction in Context*; Kurosu, M., Ed.; Springer International Publishing: Berlin/Heidelberg, Germany, 2018; pp. 241–257. [\[CrossRef\]](#)
6. Bendor, R.; Maggs, D.; Peake, R.; Robinson, J.; Williams, S. The imaginary worlds of sustainability: Observations from an interactive art installation. *Ecol. Soc.* **2017**, *22*, 17. [\[CrossRef\]](#)
7. Jacucci, G.; Wagner, M.; Wagner, I.; Giaccardi, E.; Annunziato, M.; Breyer, N.; Hansen, J.; Jo, K.; Ossevoort, S.; Perini, A.; et al. ParticipArt: Exploring participation in interactive art installations. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality—Arts, Media, and Humanities*, Seoul, Republic of Korea, 13–16 October 2010; pp. 3–10. [\[CrossRef\]](#)

8. Morrison, A.J.; Mitchell, P.; Brereton, M. The lens of ludic engagement: Evaluating participation in interactive art installations. In Proceedings of the 15th ACM International Conference on Multimedia, Augsburg, Germany, 25–29 September 2007; pp. 509–512. [\[CrossRef\]](#)
9. Pavlin, E.; Elsner, Ž.; Jagodnik, T.; Batagelj, B.; Solina, F. From illustrations to an interactive art installation. *J. Inf. Commun. Ethics Soc.* **2015**, *13*, 130–145. [\[CrossRef\]](#)
10. Bilda, Z.; Bowman, C.; Edmonds, E. Experience evaluation of interactive art: Study of GEO landscapes. In Proceedings of the 5th Australasian Conference on Interactive Entertainment, Brisbane, Australia, 3–5 December 2008; pp. 1–10. [\[CrossRef\]](#)
11. Costello, B.; Muller, L.; Amitani, S.; Edmonds, E. Understanding the experience of interactive art: Iamascope in Beta_space. In Proceedings of the Second Australasian Conference on Interactive Entertainment, Sydney, Australia, 23–25 November 2005; pp. 49–56. Available online: <https://dl.acm.org/doi/abs/10.5555/1109180.1109188> (accessed on 23 September 2026).
12. Jacquemin, C.; Gagneré, G.; Lahoz, B. Shedding light on shadow: Real-time interactive artworks based on cast shadows or silhouettes. In Proceedings of the 19th ACM International Conference on Multimedia, Scottsdale, AZ, USA, 28 November–1 December 2011; pp. 173–182. [\[CrossRef\]](#)
13. Papageorgopoulou, P.; Arsenopoulou, N.; Charitos, D.; Rizopoulos, C.; Theona, I.; Katsarou, L.; Psaltis, A.; Korosidis, A. Designing interfaces to promote the meaningfulness of urban data through an interactive art installation. In Proceedings of the Chi Greece: 1st International Conference of the ACM Greek SIGCHI Chapter, Athens, Greece, 25–27 November 2021. [\[CrossRef\]](#)
14. Yiyuan, H. Creation methodology of interactive art installation based on philosophy-understanding projection: Recreation of traditional Chinese painting. In Proceedings of the 2015 Virtual Reality International Conference, Laval, France, 8–10 April 2015. [\[CrossRef\]](#)
15. Fels, S. Intimacy and embodiment: Implications for art and technology. In Proceedings of the ACM Workshops Multimedia, New York, NY, USA, 30 October–3 November 2000; pp. 13–16. [\[CrossRef\]](#)
16. Edmonds, E. The art of interaction. *Digit. Creat.* **2010**, *21*, 257–264. [\[CrossRef\]](#)
17. Xiaobo, L.; Yuelin, L. Embodiment, interaction and experience: Aesthetic trends in interactive media arts. *Leonardo* **2014**, *47*, 166–169. [\[CrossRef\]](#)
18. Canet Sola, M.; Guljajeva, V. Visions of destruction: Exploring a potential of generative AI in interactive art. In Proceedings of the 17th International Symposium on Visual Information Communication and Interaction, Hsinchu, Taiwan, 11–13 December 2024. [\[CrossRef\]](#)
19. Guljajeva, V.; Canet Sola, M. POSTcard landscapes from Lanzarote. In Proceedings of the 14th Conference on Creativity and Cognition, Venice, Italy, 20–23 June 2022; pp. 634–636. [\[CrossRef\]](#)
20. Jacob, C.J.; Hushlak, G.; Boyd, J.E.; Nuytten, P.; Sayles, M.; Pilat, M. SwarmArt: Interactive art from swarm intelligence. *Leonardo* **2007**, *40*, 248–254. [\[CrossRef\]](#)
21. Li, J.; Sun, G.; Tang, C.; Chen, W.; Yang, W.; Kou, W.; Ruan, Z.; Ma, W.; Nie, X. Silk road journey: A real-time AI-based interactive art installation for silk road cultural reenactment and experience. In Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 11–16 May 2024; pp. 1–5. [\[CrossRef\]](#)
22. Sun, Y.; Tang, Y.; Gao, Z.; Pan, Z.; Xu, C.; Chen, Y.; Qian, K.; Wang, Z.; Braud, T.; Lee, C.H.; et al. AI Nüshu: An exploration of language emergence in sisterhood through the lens of computational linguistics. In Proceedings of the SIGGRAPH Asia 2023 Art Papers, Sydney, Australia, 12–15 December 2023. [\[CrossRef\]](#)
23. Tidemann, A.; Brandtsegg, Ø. [self.]: Realization/Art Installation/Artificial Intelligence: A Demonstration. In *Entertainment Computing—ICEC 2015*; Chorianopoulos, K., Divitini, M., Baalsrud Hauge, J., Jaccheri, L., Malaka, R., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2015; Volume 9353, pp. 517–522. [\[CrossRef\]](#)
24. Zhang, W.; Cheng, L.; Luo, J. ReCollection: Creating synthetic memories with AI in an interactive art installation. *Proc. ACM Comput. Graph. Interact. Tech.* **2024**, *7*, 1–10. [\[CrossRef\]](#)
25. Garcia, M.B. The paradox of artificial creativity: Challenges and opportunities of generative AI artistry. *Creat. Res. J.* **2025**, *37*, 755–768. [\[CrossRef\]](#)
26. Messer, U. Co-creating art with generative artificial intelligence: Implications for artworks and artists. *Comput. Hum. Behav. Artif. Hum.* **2024**, *2*, 100056. [\[CrossRef\]](#)
27. Mondal, B. Artificial intelligence: State of the art. In *Recent Trends and Advances in Artificial Intelligence and Internet of Things*; Balas, V.E., Kumar, R., Srivastava, R., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2020; pp. 389–425. [\[CrossRef\]](#)
28. Audry, S. *Art in the Age of Machine Learning*; The MIT Press: Cambridge, MA, USA, 2021. [\[CrossRef\]](#)
29. Scurto, H.; Caramiaux, B.; Bevilacqua, F. Prototyping machine learning through diffractive art practice. In Proceedings of the 2021 ACM Designing Interactive Systems Conference, Virtual Event, 28 Jun–2 July 2021; pp. 2013–2025. [\[CrossRef\]](#)
30. Plant, N.; Hilton, C.; Gillies, M.; Fiebrink, R.; Perry, P.; González Díaz, C.; Gibson, R.; Martelli, B.; Zbyszynski, M. Interactive machine learning for embodied interaction design: A tool and methodology. In Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction, Salzburg, Austria, 14–17 February 2021. [\[CrossRef\]](#)

31. Mahmud, B.; Hong, G.; Fong, B. A study of human–AI symbiosis for creative work: Recent developments and future directions in deep learning. *ACM Trans. Multimed. Comput. Commun. Appl.* **2023**, *20*, 1–21. [\[CrossRef\]](#)
32. Zhang, S.; Qi, Y.; Wu, J. Applying deep learning for style transfer in digital art: Enhancing creative expression through neural networks. *Sci. Rep.* **2025**, *15*, 11744. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Oppenlaender, J.; Johnston, H.; Silvennoinen, J.M.; Barranha, H. Artworks reimaged: Exploring human-AI co-creation through body prompting. *Proc. ACM Hum. Comput. Interact.* **2025**, *9*, 1–34. [\[CrossRef\]](#)
34. Epstein, Z.; Hertzmann, A.; Herman, L.; Mahari, R.; Frank, M.R.; Groh, M.; Schroeder, H.; Smith, A.; Akten, M.; Fjeld, J.; et al. Art and the science of generative AI. *Science* **2023**, *380*, 1110–1111. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Furtado, L.S.; Soares, J.B.; Furtado, V. A task-oriented framework for generative AI in design. *J. Creat.* **2024**, *34*, 100086. [\[CrossRef\]](#)
36. Kun, P.; Freiburger, M.A.; Løvlie, A.S.; Risi, S. GenFrame—Embedding generative AI into interactive artifacts. In Proceedings of the 2024 ACM Designing Interactive Systems Conference, Copenhagen, Denmark, 1–5 July 2024; pp. 714–727. [\[CrossRef\]](#)
37. Saltz, D.Z. The art of interaction: Interactivity, performativity, and computers. *J. Aesthet. Art Crit.* **1997**, *55*, 117–127. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Graham, C.B. A Study of Audience Relationships with Interactive Computer-Based Visual Artworks in Gallery Settings, Through Observation, Art Practice, and Curation. Ph.D. Thesis, University of Sunderland, Sunderland, UK, 1997.
39. Edmonds, E.; Turner, G.; Candy, L. Approaches to interactive art systems. In Proceedings of the 2nd International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia, Singapore, 15–18 June 2004; pp. 113–117. [\[CrossRef\]](#)
40. Chen, W.; Shidujaman, M.; Jin, J.; Ahmed, S.U. A methodological approach to create interactive art in artificial intelligence. In *Proceedings of the HCI International 2020, Late Breaking Papers: Cognition, Learning and Games: 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, 19–24 July 2020*; Stephanidis, C., Harris, D., Li, W.-C., Schmorow, D.D., Fidopiastis, C.M., Zaphiris, P., Ioannou, A., Fang, X., Sottolare, R.A., Schwarz, J., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 13–31.
41. Xie, J.; Yu, M.; Liu, G. Fusing algorithms for intersection of computer science and art: Innovations in generative art and interactive digital installations. *IEEE Access* **2024**, *12*, 173255–173267. [\[CrossRef\]](#)
42. Canet Sola, M.; Guljajeva, V. Dream Painter: Exploring creative possibilities of AI-aided speech-to-image synthesis in the interactive art context. *Proc. ACM Comput. Graph. Interact. Tech.* **2022**, *5*, 1–11. [\[CrossRef\]](#)
43. Canet Sola, M.; Guljajeva, V. Visions of destruction. In *SIGGRAPH Asia Ed. (SA '23): Special Interest Group on Computer Graphics and Interactive Techniques Conference*; Gallery, A., Ed.; ACM: New York, NY, USA, 2023. [\[CrossRef\]](#)
44. Hornecker, E.; Buur, J. Getting a grip on tangible interaction: A framework on physical space and social interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Montréal, QC, Canada, 22–27 April 2006; pp. 437–446. [\[CrossRef\]](#)
45. Benford, S.; Giannachi, G.; Koleva, B.; Rodden, T. From interaction to trajectories: Designing coherent journeys through user experiences. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, 4–9 April 2009; pp. 709–718. [\[CrossRef\]](#)
46. Sklar, S.J.; Jiang, M. Art with agency: Artificial intelligence as an interactive medium. *Humanit. Soc. Sci. Commun.* **2025**, *12*, 1546. [\[CrossRef\]](#)
47. Rombach, R.; Blattmann, A.; Lorenz, D.; Esser, P.; Ommer, B. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; Institute of Electrical and Electronics Engineers: New York, NY, USA, 2022; pp. 10674–10685. [\[CrossRef\]](#)
48. Li, J.; Li, D.; Savarese, S.; Hoi, S. BLIP-2: Bootstrapping language-image pre-training with frozen image encoders and large language models. In Proceedings of the 40th International Conference on Machine Learning, Honolulu, HI, USA, 23–29 July 2023; Volume 202, pp. 19730–19742.
49. Savaş, E.B.; Verwijmeren, T.; van Lier, R. Aesthetic experience and creativity in interactive art. *Art Percept.* **2021**, *9*, 167–198. [\[CrossRef\]](#)
50. Bilda, Z.; Candy, L.; Edmonds, E. An embodied cognition framework for interactive experience. *CoDesign* **2007**, *3*, 123–137. [\[CrossRef\]](#)
51. Candy, L. Evaluation and experience in art. In *Interactive Experience in the Digital Age: Evaluating New Art Practice*; Candy, L., Ferguson, S., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2014; pp. 25–48. [\[CrossRef\]](#)
52. Höök, K.; Sengers, P.; Andersson, G. “Sense and Sensibility”: Evaluation and interactive art. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Ft. Lauderdale, FL, USA, 5–10 April 2003; pp. 241–248. [\[CrossRef\]](#)
53. Szubielska, M.; Imbir, K.; Szymańska, A. The influence of the physical context and knowledge of artworks on the aesthetic experience of interactive installations. *Curr. Psychol.* **2021**, *40*, 3702–3715. [\[CrossRef\]](#)

54. Zhou, C.; Li, J. The development of aesthetic experience through virtual and augmented reality. *Sci. Rep.* **2024**, *14*, 4290. [[CrossRef](#)] [[PubMed](#)]
55. Heid, K. Aesthetic development: A cognitive experience. *Art Educ.* **2005**, *58*, 48–53. [[CrossRef](#)]
56. Joy, A.; Sherry, J.F., Jr. Speaking of art as embodied imagination: A multisensory approach to understanding aesthetic experience. *J. Consum. Res.* **2003**, *30*, 259–282. [[CrossRef](#)] [[PubMed](#)]
57. Leder, H.; Belke, B.; Oeberst, A.; Augustin, D. A model of aesthetic appreciation and aesthetic judgments. *Br. J. Psychol.* **2004**, *95*, 489–508. [[CrossRef](#)] [[PubMed](#)]
58. Caramiaux, B.; Fdili Alaoui, S. “Explorers of unknown planets”: Practices and politics of artificial intelligence in visual arts. *Proc. ACM Hum. Comput. Interact.* **2022**, *6*, 1–24. [[CrossRef](#)]
59. Allen, S. *Unreal Pareidolia -shadows-*. Scott Allen. 2023. Available online: <https://scottallen.ws/work/unreal-pareidolia-shadows> (accessed on 23 September 2026).
60. Allen, S. *Unreal Pareidolia -shadows-*. CVPR AI Art Gallery. 2024. Available online: <https://thecvf-art.com/project.php?year=2024&artist=scott-allen&id=786> (accessed on 23 September 2026).
61. Zhang, W.; Luo, J. Re:collection. In Proceedings of the SIGGRAPH 2023 Art Gallery (SIGGRAPH ‘23): Special Interest Group on Computer Graphics and Interactive Techniques Conference, Los Angeles, CA, USA, 6–10 August 2023. [[CrossRef](#)]
62. Zhang, W.; Luo, R. *ReCollection*. CVPR AI Art Gallery. 2024. Available online: <https://thecvf-art.com/project.php?year=2024&artist=recollection&id=687> (accessed on 23 September 2026).
63. Guljajeva, V.; Canet Sola, M. *Visions of Destruction*. CVPR AI Art Gallery. 2024. Available online: <https://thecvf-art.com/project.php?year=2024&artist=varvara-mar&id=758> (accessed on 23 September 2026).
64. Jeon, M.; Fiebrink, R.; Edmonds, E.A.; Herath, D. From rituals to magic: Interactive art and HCI of the past, present, and future. *Int. J. Hum. Comput. Stud.* **2019**, *131*, 108–119. [[CrossRef](#)]
65. Radford, A.; Kim, J.W.; Xu, T.; Brockman, G.; McLeavey, C.; Sutskever, I. Robust speech recognition via large-scale weak supervision. In Proceedings of the International Conference on Machine Learning, Honolulu, HI, USA, 23–29 July 2023; Volume 202, pp. 28492–28518.
66. OpenAI. GPT-4 Technical Report. *arXiv* **2023**, arXiv:2303.08774.
67. Wiberg, M.; Stolterman, E. Time and temporality in HCI research. *Interact. Comput.* **2021**, *33*, 250–270. [[CrossRef](#)]
68. Bruckner, D.; Velik, R.; Penya, Y.K. Machine perception in automation: A call to arms. *EURASIP J. Embed. Syst.* **2011**, *2011*, 608423. [[CrossRef](#)]
69. Nevatia, R. *Machine Perception*; Prentice Hall: Hoboken, NJ, USA, 1982.
70. Guile, D.; Popov, J. Machine learning and human learning: A socio-cultural and -material perspective on their relationship and the implications for researching working and learning. *AI Soc.* **2025**, *40*, 325–338. [[CrossRef](#)]
71. Hutchins, E. *Cognition in the Wild*; The MIT Press: Cambridge, MA, USA, 1995. [[CrossRef](#)]
72. Walker, R.J. *Seeing Ourselves Through Technology How We Use Selfies, Blogs and Wearable Devices to See and Shape Ourselves*; Springer Nature: Berlin/Heidelberg, Germany, 2014.
73. Shaw, J. The legible city. In *The Boundary Rider: 9th Biennale of Sydney*; Bond, A., Ed.; Biennale of Sydney: Sydney, Australia, 1992; pp. 387–398. Available online: https://www.jeffreyshawcompendium.com/wp-content/uploads/2018/03/2001_Im-Buchstabenfeld.EXT_GERMAN.pdf (accessed on 23 September 2026).
74. Peck, E.J. Understanding Digital and Interactive Media in Live Events and Performance: The RAIN ROOM. Academia. 2018. Available online: <https://www.academia.edu/download/57279835/RainRoom.pdf> (accessed on 23 September 2026).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.