UC Riverside
Conference Papers

Title
Maintained by Omission

Permalink
https://escholarship.org/uc/item/61h0x0pn

Author
Mares, Ariana

Publication Date
2026-06-01

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Maintained by Omission

Ariana Mares
University of California, Riverside
amare029@ucr.edu

Abstract

Maintained by Omission is an audiovisual composition that explores the tension between habitual human perception and machine-assisted perception. Rooted in Vilém Flusser's media philosophy, the project approaches artificial intelligence not as a conscious entity, but as an interactive processing mirror within a cybernetic, dialogic feedback loop. Grounded in the premise that human cognition routinely filters out environmental noise to maintain perceptual stability and social balance, keeping us from being overwhelmed by the world's constant, unfiltered saturation, this paper takes the reader on an inward journey through the project's creative process—mirroring how the composition itself turns inward toward the hidden layers of human perception and daily life. Through an interactive dialogue across language models (ChatGPT, Gemini), audio architectures (Max/MSP Somax 2), generative video models (RunwayML), and AI audio generators (Suno), raw human observations—captured via environmental field recordings, violin improvisations, spoken poetry, and historical audio archives—are calculated, shuffled, recombined, and recomposed. By embracing play and chance over strict control, the project shifts the creative process from a discrete 'chess game' toward an embodied, continuous dance with the machine. Ultimately, *Maintained by Omission* demonstrates how human-machine co-creation can expose the hidden architectures of daily perception, transforming omitted noise into a rich, shared space of intersubjective synthesis.

Keywords: Audiovisual composition; Human-AI co-creation; Vilém Flusser; Interactive multi-agent systems; Sensory perception

1. Introduction

What would happen if we stopped our mundane routine? What would we notice for a single minute if we gave ourselves the time to see? What would you see? Is it empty? Is it filled with little moments we call accidents? What does it reflect about who you are? Who we are?

Observing, listening, detecting, recognizing, remembering, and perceiving are actions we do not allow ourselves to take once in a while, for whichever valid reason we may have?

These are some of the ideas that floated through my mind when presented with the challenge to create an audiovisual composition using AI in Spring 2026 during MUS249 Audiovisual and Multimedia Composition taught by Dr. Paulo C. Chagas at UCR.

How am I supposed to incorporate such a human experience into a creative outlet where I have to collaborate with a machine such as Artificial Intelligence? Could I rely on a machine/apparatus? Could AI comprehend the multi-faceted nature of human life and perception? Sure, I wanted to explore the idea that stopping and observing our surroundings in the moment helps us to realize that much more is happening around us than we may think, but exploring these questions would also require examining how stopping to observe our

(immediate) surroundings could reveal the hyper-present horizon of latent activity that human cognition routinely filters out in order to maintain a stable reality.

Instead of expecting the AI systems to possess comprehension or subjective consciousness, my project approaches the machines as computational processing agents within a cybernetic feedback loop. The AI apparatuses cannot "feel" or "perceive"; instead, they calculate probabilities across data streams within the context I provide to them. Through a continuous dialogue with the apparatus, my own raw human observation—captured through environmental field recordings, vocal recitations, and violin improvisations—is then fed into multi-agent systems (Max/MSP Somax 2) and generative models (Suno, RunwayML). Ultimately, the machine acts as an interactive processing mirror: it calculates algorithms and recombines sound and image, creating improbable layers of daily noise that human sensory habituation routinely omits.

This kind of interaction redefines human-machine collaboration and participation. In *Into the Universe of Technical Images*, Vilém Flusser compares telematic dialogue to a chess game—a discrete, turn-based contest governed by abstract rules and formal strategy (Flusser, 2011, Ch. 11–12). However, the real-time audiovisual synthesis used in *Maintained by Omission* differs from Flusser's chess game in crucial ways. While a chess game, according to Flusser, operates through discrete, symbolic moves on a grid, my work with generative AI instead functions as an embodied, continuous dance and improvisational chamber-music ensemble (Flusser, 2011, Ch. 19).

Flusser rejects the idea of creation as a miracle occurring from nothingness. Instead, he argues that creation emerges from two main variables: information and interaction. For Flusser, information arises from synthesizing previously acquired information. In other words, new data is created by accumulating and assimilating previous data and information already available to us—or, as he describes it, "people are players with prior information" (Flusser, 2011, p. 89).

One important and essential aspect of Vilém Flusser's media philosophy is the mechanism through which human culture generates new information. Flusser distinguishes between natural processes and human play. The former operates randomly and without a purpose; the latter, he says, is purposeful: "Nature plays without purpose... human beings play using dialogue" (Flusser, 1985/2011, p. 90). For Flusser, dialogue is not what we call and immerse ourselves daily in "casual conversation," but instead, the intersubjective exchange and synthesis of stored data across different memories—both human and artificial—to produce unpredictable, novel information that counteracts entropic decay (Flusser, 2011, pp. 83, 89). Through this framework, Flusser dismantles the Romantic ideal we have built as a society of the isolated artist, the one who creates out of nothing through purely internal reflection. Instead, he argues that creation in the era of technical media is a recombinatory game, where new information is produced by calculating, shuffling, and synthesizing pre-existing data rather than creating out of nothing (creatio ex nihilo) (Flusser, 2011, pp. 89, 116). To illustrate this, Flusser compares outer dialogue to a chess game played by multiple participants. Once a participant—whether another human or a technological apparatus—joins the game, the possibilities for generating new, unexpected information expand exponentially (Flusser, 2011, pp. 100–101).

As Flusser mentions, these interactions and outer dialogues can occur between human beings and the apparatus. The interaction we have with the apparatus, in this case AI, produces new material/information, but only through sharing information and dialogue between human beings and the apparatus/machine. The interplay of ideas and data from both ends creates a playful combination of collective information. In this sense, the new creation is a combined result of synthesized information.

Building on this idea, Flusser critiques the traditional notion of the "I"—the idea of the individual as a self-contained cell that must preserve its core from external influence. Drawing on depth psychology, informatics, and phenomenological analysis, Flusser dismantles this isolated atomism, arguing that an individual's singularity is not preserved in solitude but revealed through interdependent connection with others and technological apparatuses; the "I" exists precisely because we exist in relation to others. It is revealed through the freedom of socialization with others; the "I" (individual) exists because we are with others (Flusser, 2011, pp. 120–123).

With this idea in mind, this shift from individuality toward collective collaboration is not entirely new, as we have been interacting with the other and the apparatus from the very beginning. For example, a photographer and their photographic apparatus make creative play a necessary interaction for mutual recognition. In this sense, creative production becomes a collective effort between human artistic intention and the apparatus's program.

2. Project Concept and Documentation of Process

The contrast between the usual lack of awareness we have as human beings of what happens around them/us, and the way a person with enhanced senses, such as autism, perceives the world, has always fascinated me. If we could meet somewhere in the middle and learn from each other's world, we would notice how much we can learn to see our day-to-day life with a completely different perspective. We face the same challenge in our relationship with technology, and even more so with AI today: we often struggle to step out of our automated routines and habitual perceptual filtering to engage with a way of sensing and processing reality that is, in a way, radically different from our own. With this idea in mind, and knowing that a conversation between the individual—in this case, myself—and various apparatuses such as Gemini and ChatGPT would be necessary, I opted to start with ChatGPT for its strengths in writing, creativity, and even its hallucinatory responses at times. In my conversations with ChatGPT, it gave me ideas on what to capture before even starting the project:

> "Because this is a co-creative AI project, the material should be collected in ways that give AI *ambiguity* to work with. AI becomes strongest when the source material is rich, textured, unstable, repetitive, or partially unreadable. Collect recordings in *layers*, not isolated "clean" sounds. You do *not* want perfectly composed nature photography. You want material AI can destabilize." –*ChatGPT*

Figure 1. Response from ChatGPT towards the main idea for the project

After making my own interpretations of the words and ideas ChatGPT gave me: "layers," "clean sounds," "rich, textured, unstable, repetitive, unreadable," I collected all the footage I deemed appropriate (and, once more, to my interpretation) was necessary; and this time, in conversation with Gemini, it helped me create a conceptual idea for the video. It is important

to clarify that the same conversations would happen with ChatGPT and Gemini, noticing that each AI has its strengths and weaknesses depending on the questions asked. In this case, Gemini connected my ideas to those expressed at the beginning.

The Concept: *The everyday environment (your field recordings and footage) seems boring and repetitive because we are "programmed" to filter it out.*

The Co-Creative Act: *By using AI to isolate, amplify, and recombine these ordinary sounds and images, you are playing a "game" against human sensory limitation. You are forcing the latent, improbable layers of reality to the surface, transforming noise into new Information.*

-Gemini

Figure 2. Conceptual idea from Gemini

With this new information in mind, I was reminded of a scene from the movie The Peaceful Warrior (2006), directed by Victor Salva and adapted from the book "Way of the Peaceful Warrior" by Dan Millman. In it, the mentor (Socrates) shows the athlete (Dan) that when we pay attention, we notice that "there is always something happening around us." Building on that idea, and wanting to weave in a few different theoretical concepts, from Katherine Hayles' Semiotic Square of the Posthuman and Michael Chion's listening modes to Nicholas Cook's models of multimedia relationships, I asked ChatGPT to help me organize it:

That direction actually fits your material *extremely* well.
And it shifts the piece from:

- memory collapse,
- haunted environments,
- unstable archives,

toward something more: **phenomenological and spiritual.**
Not spirituality in a religious sense necessarily, but heightened awareness of invisible activity. The scene from Peaceful Warrior works because the revelation is not: "something magical appears." It is: the world was already alive the entire time.

-ChatGPT

Figure 3. Response from ChatGPT when organizing concepts studied in class

Adding to this dialogue between ChatGPT and Gemini, I knew the only way (at least for me at the time) the project would work was to incorporate text in the form of a poem. My decision to use Gemini instead of ChatGPT came from previous experiences with ChatGPT's generated poetry and from seeing the experiences of others with each apparatus.

The Co-Created Poem: *The Lattice of the Present*

(Designed to be delivered as a sparse, low-passed voiceover, or projected as minimalist technical text overlays that fragment and dissolve over the 8-minute duration).

I. The Filtered Surface we walk through a landscape of heavy curtains, trimming the edges of the wind to keep the quiet. We call it silence when the machine is still, Predicting the horizon, tracing a pattern we already expect. The mind draws a circle around the tree and calls it empty.

II. The Fracture But the shadow is not an absence; It is a weight. A friction of bark against bark, A microscopic fracture in the air. Listen to the error. Listen to the sharp, accidental seam where the template breaks. The apparatus remembers what you threw away.

III. The Horizon there is never nothing happening. The canopy is an intricate, mathematical score, A cascade of simultaneous architectures. Every droplet a pulse, Every decay a new velocity. The noise is not chaos—it is crowded with presence. We are unmasked by the texture. We are awake in the lattice.

-Gemini

Figure 4. First poem created with Gemini

For the concept I was co-creating with both Gemini and ChatGPT, the poem felt, in this case, too explicit and, through my creative lens, did not leave enough room for multiple interpretations for the audience. Similarly, I wanted to create an abstract audiovisual composition, leaving room for AI hallucinations when it started generating the video scenes. As a result of this problem, I turned to Gemini:

> *"I would like the poem to be more abstract, that the viewer can truly internalize it and connect with it from their own experiences and perceptions. In that way, the poem, audio, and visuals become intricately unique to the individual that is listening"*

Gemini generated the poem that was ultimately used for the video:

The Co-Created Poem: An Architecture of Gaps

I. It is maintained by omission.
The long, quiet line
we draw around the day,
predicting the horizon,
assuming the frame is full.
You sit within the rhythm
of an expected room,
filtering the hum,
calling the stillness empty.

II. But the outline holds a weight.
A friction at the edge of the screen,
where the template fractures,
and the memory slips.
Listen to the accident.
The sharp, unchosen seam
where the expectation breaks,
and the archive begins to breathe.

III. There is no interval.
The space is a crowded geometry,
a cascade of simultaneous returns.
Every shadow a body,
Every fracture a sudden velocity.

You are unmasked by the texture.
You are awake in the lattice.
Nothing is waiting to begin.

-Gemini

Figure 5. Final poem created with Gemini

For the music, my first instinct was to avoid using Suno, as generative music platforms often default to overly polished, commercial aesthetic outcomes. Because Suno in itself was designed by commercial developers and trained on massive datasets of mainstream pop and commercial recordings, its algorithms naturally lean toward conventional song structures rather than abstract, electroacoustic soundscapes. Such problem, for me, occurs due to the purpose and context for which Suno was created; in my case and creative purpose, as a default, it would not work. Wanting to preserve the composition's experimental, abstract core, I initially hesitated. However, curious about the playfulness and dialogic interplay between human and machine, I first had Gemini create the music, but to my dissatisfaction, it created either pop or hip-hop songs: Leaves on the Concrete. Because of this, I decided to use Gemini as a prompting mediator to craft highly specific text prompts that could push Suno beyond its commercial training defaults—though this iteration still produced unsatisfactory results. The prompts that ended up being useful in creating the four songs/pieces of music used in the final video were:

Cinematic ambient, experiential drone, experimental soundscape, slow tape hiss, granular synthesis, whispered spoken word, subterranean low-pass filter, microtonal strings, sparse piano resonance,
60 BPM: Omission Horizon_short , Omission Horizon_longer

Cinematic ambient, drone soundscape, slow glacial chords, microtonal texture, deep background whispered spoken word, ethereal voice projection, vast reverb, non-rhythmic, no drums, no percussion, abstract sound design, infinite decay,
55 BPM: Long version , Shorter version

Figure 6. Prompts generated by Gemini to be used in Suno

Because of the complexity of the collaboration, the project began to unfold from the interplay between human perception (mine) and the ideas created by the dialogue between ChatGPT, Gemini, and me; the obvious next step was to take all the musical elements into the Max MSP program, Somax. Having in mind the previous ideas from ChatGPT about phenomenology and the idea from Gemini about playing with the human sensory limitations to create new information, the next course unfolded by itself:

How would four players in Somax's Interactive platform react to my voice reciting the poetry? How would they react to the songs SUNO created?

For this process, I thought adding a human element was important. The four interactive players, while in themselves are AI-generated, the material they were going to use was going to be pre-existing recordings as well as recordings of my own. The four players interchanged material from the recordings and field recordings, ensuring cohesion. I chose the human aspect: which recordings to use and which songs to use from Suno, but even those decisions came from our dialogue (Individual vs. Apparatus). During my session with Somax, I chose which audio to use at the moment, but which corpora would interact with one another was more of a matter of chance. While I could not predict the outcome of the interaction, I would, from time to time, choose which corpus to bring out more and when to change the audio. Similarly, once I collected all the audio from both Somax and Suno, I chose the editing, but again, it was through an intangible dialogue between the audio results and me. They were later mixed in Logic Pro, adding new layers to each audio and following an arc guided by the music generated on Suno.

Audios used for the LogicPro Session: Audios

Lastly, for the videos, the poem created by Gemini was used to generate each video, using each line for the creation of the video. The initial idea was to use prompts from Gemini:

In conversation with Gemini: Give me prompts for videos that will help me convey the poem

Text Prompt: > Cinematic extreme slow-motion, subtle atmospheric shift, a microscopic breeze faintly moving a single leaf, hyper-stable camera, maintaining perfect structural geometry of the shadow, high-contrast black and white, 35mm film grain, absolute tranquility.

Result: https://app.runwayml.com/creation/090c6350-f481-4a76-9fb6-fb81c4044280

Figure 7. Prompts generated by Gemini to be used in Runway. The poetry was used to model the prompts.



Figure 8. Images generated by Runway from Gemini's Figure 7 prompt.

The videos, while somewhat abstract, did not convince me of the project's narrative and concept. Instead, the poem itself was used to generate the videos, along with some videos I prompted with the concept of the audiovisual composition in mind.

Prompt: *I want something creative and abstract where I am playing with people's perspective. I have multi-layered poetry that is making people aware of everything that is happening around them. In the audio there is also a fountain of water and bird calls.*

Result1: https://app.runwayml.com/creation/6c61900a-4d8f-4c2b-82e3-648326a0f6ca
Result 2: https://app.runwayml.com/creation/88713b7d-5029-4e98-82e9-69dd5fff642b



Figure 9. Images generated by Runway using my own prompt, based on the project's main concept.

Prompt: *A dark room with a very small entry of light, the circle/crack starts becoming biggerand bigger*

Result: https://app.runwayml.com/creation/c891dcf3-0adf-4869-8338-b654969647df



Figure 10. Image generated by Runway using my own descriptive prompt.

Prompt: *There is silence and darkness at the beginning, and we hear four gasps simultaneously, and the poem starts: It is maintained by omission. The long, quiet line we draw around the day, predicting the horizon, assuming the frame is full.*

Result: https://app.runwayml.com/creation/0d52bd9e-cbab-4ae5-bc87-c0b80749e6e4



Figure 11. Image generated by Runway using my own descriptive prompt.

Some videos were the exact same ones Runway generated for the poem, but others I chose for their imagery and audible action. For example, the following prompt generated a couple of quick scenes that were used at different moments throughout the video. These moments were outside of the poetry line they were generated with:

Prompt: *Next part of the poem: But the outline holds a weight. A friction at the edge of the screen, where the template fractures, and the memory slips.*

Images with the moment they were used instead:



Used with a black-and-white filter and flickering at the beginning with the second breath and no words.
Used again with text: the long quiet line

Used with text: we draw around the day





Used with text: predicting the horizon

Used with text: predicting the horizon



It is important to note that the audio was edited first, before even starting to generate the videos. In the same way, not all the footage used in the final project is AI-generated; some videos and images are footage I took with my phone, and I added filters if necessary to convey the message of the audio and the overall concept.

The AI-generated content, whether co-created or given the freedom to generate and hallucinate, completely changed and shaped the overall audiovisual composition. The most important change was embracing abstraction as the common thread uniting the image and the sound. Given that I was not sure as to how I was going to put both elements together, the playfulness of chance and text-driven choices informed the pace of the piece.

Another important aspect was the relationship between sound and image, which involved the use of filters in the video editing through Filmora –a non-linear video editing application designed for timeline-based video compositing, multi-layering, and audio-visual synchronization. The decisions about effects and minimal transitions were driven by the text and the result of the poem treatment through Somax. For example, in the section (5'10'') where in the audio we hear the chaotic interaction between Gabriela Mistral's poem *Canción Quechua*, Bartók's *Presto* for violin solo, and improvisations on violin, the images and videos start to change rapidly. The decision came through the transformation of the primary audio, affecting how the image interacts with the sound.

As to why the usage of audios such as Béla Bartók's *Presto* for violin solo and Mistral's poem *Canción Quechua*, the latter, Gabriela Mistral's poem from a Chilean poet, encapsulates not only my Chilean heritage but also the cosmovision of the Pre-Columbian cosmovision, in which the idea of awareness of the individual, the other, and the environment is embodied in their day-to-day life, human philosophy, and way of living. For the former, Béla Bartók, a Hungarian composer and pianist, was known for the use of folk tunes embedded in his music.

Some definitions to understand Gabriela Mistral's poem

Tahuantinsuyo: Pre-Columbian Inca Empire

Vicune: a wild animal similar to a llama; it inhabits the mountains in South America

Pachacamac: large pre-Columbian ruin located in the Lurin Valley on the central coast of Peru

Figure 12. Mistral's poem word definitions.

CANCIÓN QUECHUA

Donde fue Tihuantisuyo,
nacían los indios.
Llegábamos a la puna
con danzas, con himnos.

Silbaban quenas, ardían
dos mil fuegos vivos.
Cantaban Coyas de oro
y Amautas benditos.

Bajaste ciego de soles,
volando dormido,
para hallar viudos los aires
de llama y de indio.

Y donde eran maizales
ver subir el trigo
y en lugar de las vicuñas
topar los novillos.

¡Regresa a tu Pachacamac,
En-Vano-Venido,
Indio loco, Indio que nace,
pájaro perdido!

QUECHUA SONG

In the area where Tihuantisuyo was,
the American Indians were born.
We reached a plateau in the Peruvian Andes
with dances, with hymns

They were dancing the Inca women and
live fires for two thousand years.
They were playing the golden flutes
and blessing the Inca blood.

You came down blind from the suns,
falling asleep, and when you woke up
you found the widows and the turtledoves
crying in the Indian wind.

And where used to be cornfields
now you see the wheat rising
and instead of vicune
you are bumping into cowboys and cattle.

Return to your Pachacamac,
You come in vain,
crazy Indian, you Indian boy,
the lost bird!

Gabriela Mistral
(1889-1957)

Figure 13. Poem *Canción Quecha* by Mistral, Spanish and English translation.

Some, but not all, of the lines are heard through the audio of Maintained by Omission. Both Bartók and Mistral were used in the interactive four players in Somax. If one pays close attention, some of the material is recognizable.

.

3. Reflections

The AI-generated content, whether co-created or given the freedom to generate and hallucinate, completely changed and shaped the overall audiovisual composition. The most important change was embracing abstraction as the common thread uniting the image and the sound. Given that I was not sure as to how I was going to put both elements together, the playfulness of chance and text-driven choices informed the pace of the piece.

Another important aspect was the relationship between sound and image, which involved the use of filters in the video editing through Filmora –a non-linear video editing application designed for timeline-based video compositing, multi-layering, and audio-visual synchronization. The decisions about effects and minimal transitions were driven by the text and the result of the poem treatment through Somax. For example, in the section (5'10'') where in the audio we hear the chaotic interaction between Gabriela Mistral's poem *Canción Quechua*, Bartók's *Presto* for violin solo, and improvisations on violin, the images and videos start to change rapidly. The decision came through the transformation of the primary audio, affecting how the image interacts with the sound.

Seeing how the material would transform through a system built by the apparatus was probably the most rewarding outcome. While I faced the obstacle of video generation from Gemini's prompts, deciding to use the lines of the poem and trusting what Runway would generate from them allowed the interplay, collaboration, release of control, and a creative dialogue between individual and machine. The results of the videos generated by Runway are what gave the core of the audiovisual composition itself.

From the very beginning, I wanted to be in charge of the project; I wanted to decide what to pay attention to, what to photograph and film, and how the music should be. When I found myself fighting with the apparatus, I would give up and start a conversation with AI; that way, I could be playful with the content. In the same way, the dialogue and playfulness would tend to be interchangeable and feed one another. The resistance I felt shows the internal fight we have with a new experience, which appears only when we interact with the apparatus, or, as Flusser says, when the "I" is revealed through the other. My project shows the possibilities once one abandons prejudices against co-creation through free dialogue and playfulness with the other.

4. List of tools and AI systems used

4.1. AI Systems

ChatGPT
Gemini
Runway
Somax
Suno

4.2. Audio and Video Editors

Logic Pro
Filmora (video editor)

4.3. Content

Footage (taken on iPhone 16 Pro)

Field Recordings

- Birds
- Water fountain
- Train
- Motorcycle
- Cat meowings
- Male voice
- Rustling leaves
- Tower Bells

Recordings

- Bartók for solo violin (Presto)
- Gabriela Mistral reciting her poem “Canción Quechua”
- Improvisations on Violin
- Improvisations on Voice (non-singing)

5. Final Video

https://audiovisualmusic.ucr.edu/co-creativity-exhibition/maintained-by-omission/

6. References

Chion, M. (1994). *Audio-vision: Sound on screen* (C. Gorbman, Trans.). Columbia University Press.

Cook, N. (1998). *Analysing musical multimedia*. Oxford University Press.

Flusser, V. (2011). *Into the universe of technical images* (N. A. Roth, Trans.). University of Minnesota Press. (Original work published 1985)

Hayles, N. K. (1999). *How we became posthuman: Virtual bodies in cybernetics, literature, and informatic*s. University of Chicago Press.

Jung, C. G. (1959). *Aion: Researches into the phenomenology of the self* (R. F. C. Hull, Trans.; Collected Works of C. G. Jung, Vol. 9, Part II). Princeton University Press. (Original work published 1951)