Concordia University

FINAL REPORT

JAMAIS-VU:
Archiving Urban Soundscapes, the Recorded Ghosts of Our Present and a Potential Future

Lydia, Azhar, Louis, Noah, Kamyar, Rebecca

Documentation Website

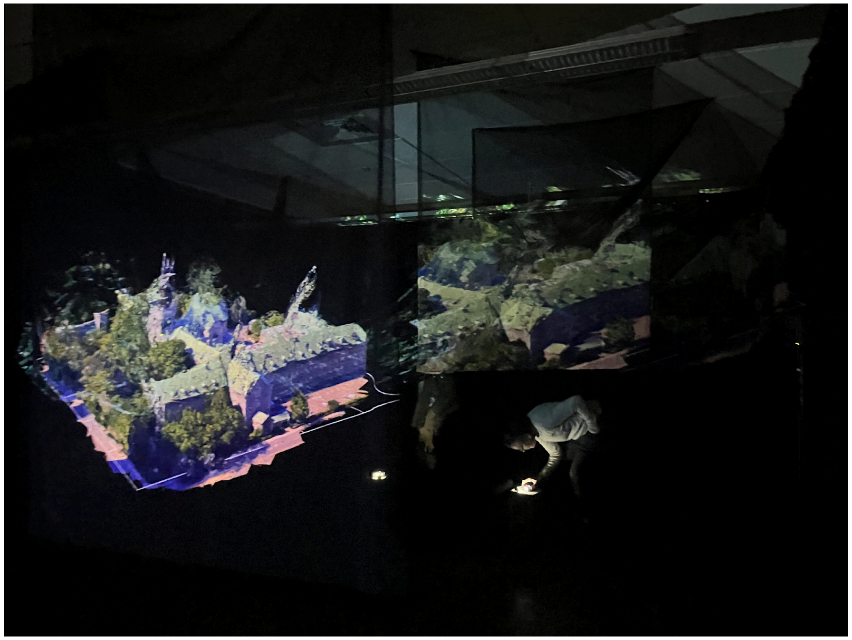
CART 461 - TANGIBLE MEDIA STUDIO

Elio Bidinost

December 1st, 2024

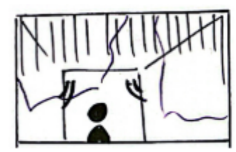
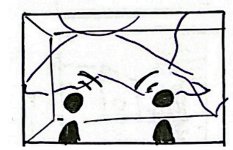
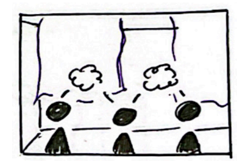
Abstract

Jamais Vu is an immersive synaesthetic installation that transforms a gallery space into a dynamic multisensory environment. Our project takes its name "Jamais Vu" from the phenomenon of experiencing unfamiliarity when encountering something that should be familiar. The project seeks to disrupt habitual familiarity with our increasingly highly digitized world, in which more and more aspects of our lives are being tracked, measured, and converted into data, reducing people from complex individuals into quantifiable units of information. We aimed to evoke this sensation by creating a (syn)aesthetic installation featuring a soundscape composed of crowdsourced audio samples from across the city of Montreal. This defamiliarization may prompt an understanding of how the present actively intrudes upon, reshapes, and haunts our possible imaginations of the future.



Design Narrative

The installation features a central projection area of suspended fabric sheets onto which audio-reactive visuals are projected, representing a digitized portrayal of Montreal. Participants are invited to wander throughout and interact with the space from multiple angles. The room is enveloped by a spatialized generative soundscape, made by our participatory archive of crowdsourced audio submitted to our website. These sounds—everyday recordings of the city—are passed through a machine learning process that fragments, distorts, and reassembles these sounds into a reimagined algorithmic city. Handheld recording devices placed throughout the installation allow participants to contribute new sounds in real-time, providing a tactile way to engage while inside the installation. This interplay of different mediums and modes of participation is designed to create a space for interpretation and subjective sensory experience that resists being reduced to a singular, data-driven narrative. Each iteration of Jamais Vu is unique, consisting of a new combination of sounds from the past and present, always changing and reacting to new data fed by participants and giving agency back to the audience through the data collection process. In this way, the installation becomes a kind of haunted space: a simulated environment in which the boundaries between real and unreal, past and present, are fluid.

ENTRY / DISCOVERY 	ENGAGE / EXPLORE 	REFLECT 
<u>First impressions</u> Greeted by some ambient sound coming from the dark room with visuals are projecting onto fabric sheets <u>Participants are curious and intrigued</u>	<u>Interaction with the space</u> They navigate and explore the space, engaging with the recording device and watching visuals respond dynamically to the sounds of the city <u>Participants are captivated</u>	<u>Reflection zone</u> They take a seat in the space to reflect and may also discuss with other participants their thoughts on the experience <u>Participants are immersed, inspired and connected</u>

Prototype Process

The project's initial intention has shifted over the course of researching and implementing the physical prototype, particularly in terms of the scale and method of interaction. Originally, the project aimed to create an expansive and city-wide experience where participants could contribute audio samples through a web app or physical devices distributed across Montreal. The intention was to evoke "jamais vu" by collecting diverse sounds from the city's urban environment, creating a participatory and autopoietic soundscape shaped by external contributions and empowering participants by giving them agency in data collection.

We had to resolve the tension between the multiple options we had for data collection: the app and the physical device. We had to ask ourselves: is this device being a part of exterior data collection going to amplify the mission of the installation? What story are we trying to tell with the interactions and the installation? Rather than distributing recording devices for collecting sounds across the city, we decided to focus more on the installation as a site for meaningful interaction rather than solely a presentation space, where the interaction would be localized within the exhibition. Our revised approach encourages engagement by allowing participants to record sounds within the installation using strategically placed devices throughout the space, while the main data collection for the urban soundscape will be done via the web application, which will be shared using QR codes. This setup not only fosters interaction but also invites viewers to explore and immerse themselves fully in the environment. The audience using the device inside the installation allows for an embodied immersion of what would be our roles vis-a-vis contribution to unimagined futures, generated by machine learning models.

Additionally, instead of relying solely on crowdsourced contributions to build the soundscape, we adapted our approach by preparing a baseline of sounds beforehand. This adjustment was necessary to ensure that the installation had enough initial material to be functional and engaging from the start. The shift doesn't deviate from our core intention of creating a sensory, defamiliarizing experience but does adjust the method of achieving it to a more focused, controlled environment.

Machine Learning and Computing Power

Our intention with machine learning remained as it was, however we had to process different Machine Learning algorithms and also softwares that allow us to achieve it in order to achieve the final result. We initially wanted to use RAVE, A variational autoencoder for fast and high-quality neural audio synthesis, to train a model fed on our field recording data. However, this option seemed infeasible given the computing power that we had access to. Training our own model would have taken a lot of time and we had to look for more scoped options. We ended up using AudioStellar, which is a machine learning powered sampler developed by Instituto de Investigaciones en Arte y Cultura "Dr. Norberto Griffa". AudioStellar gave us the possibility to use our own small dataset to recreate urban sounds without the need for a super computing power. AudioStellar's built in K-Clustering algorithm presented newer ways of looking at sound data, and we were able to create mutated and randomized sound sequences to recreate urban sounds from our dataset that weren't necessarily related but rather gave the illusion, which to our aim was aligned with the feeling of “Jamais-Vu”.

INSTALLATION DESIGN

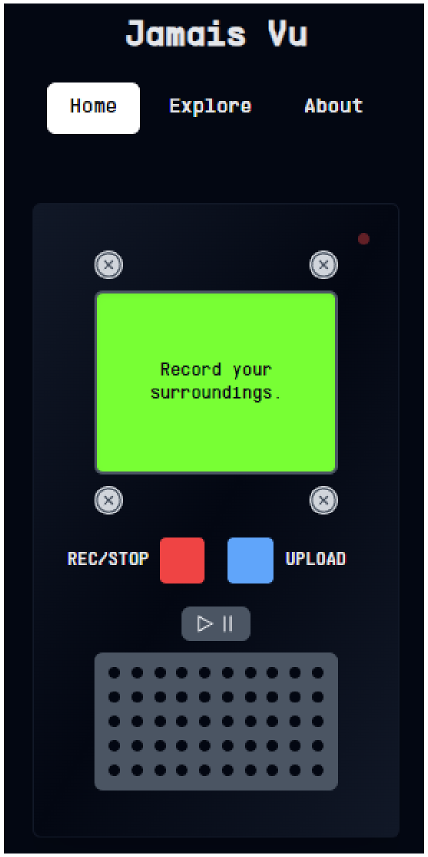
Given how site-specific the installation can be, we reiterated our layout design multiple times to achieve a welcoming and explorable experience specific to the holodeck. We decorated the room mostly with sheets of black translucent fabric and this combination with projected and mapped visuals gave way for a simulated urban space. At the final iteration, we decorated the room in a way that resembled avenues and streets so that the audience could move forward to the end (resembling traffic). At the end of the room we placed a bench with two concretes to the sides each holding a light and the device to welcome the audience to interact with the space and the devices.

DEVICE DESIGN/DEVELOPMENT

For the recording device, we stuck with the idea of having it in the installation rather than being used to collect audio samples around the city. This change was made to add an extra layer to the installation but also because we already had the web application to record sounds.We decided to use two particle Argons over a local network to record sound within the installation. This meant that we were able to send WAV data to a Node.js server. While the device recording was great, we wanted to make the device more inviting for users to interact with while in the installation. We ended up adding an OLED screen which contained prompts and leds which would change based on the state of the device.
As for the design of the device, our intention was to keep it a small hand-held to record voice, including a display and a button to facilitate user interaction.
The final shell was built using 3D-printed black resin which was sanded and varnished to have a translucent finish, exposing the inner components of the device.
On the inside, each device has a prototyping board that houses all the electronic components: a PDM mic, an OLED display, a button, a battery, and a Particle Argon. To set it up in the space, we placed the two devices on concrete blocks with a small light shining onto them. They were placed on each side of a small bench, which we added into the space to encourage visitors to sit and interact with them.

APP DESIGN/DEVELOPMENT

The app maintained its primary function of collecting audio recordings, but received a visual update to align with the device aesthetics. A heatmap feature was added, enabling users to see where audio files were recorded and play them back. During testing, it was discovered that some users' geolocation data failed to reach the server.



To resolve this, the user's IP is now used as a fallback to approximate their location. The app, hosted on Vercel, was built using the T3 Stack, which includes TypeScript, Next.js, Tailwind CSS, and tRPC. IBM Cloudant is used for the database to store geolocation data, while IBM Cloud Object Storage handles the storage of audio files.

Other NPM packages used:

➔ ShadCN UI: Used for the "Navigation Menu" component.

➔ Wavesurfer.js: Displays the waveform of audio recordings.

➔ MapTiler: Generates the heatmap feature.

➔ Radix Icons: Provides icons like "play" and "pause."

➔ Extendable-media-recorder: Records and encodes audio files in WAV format.

➔ GeoJSON: Formats geolocation data for consistency.

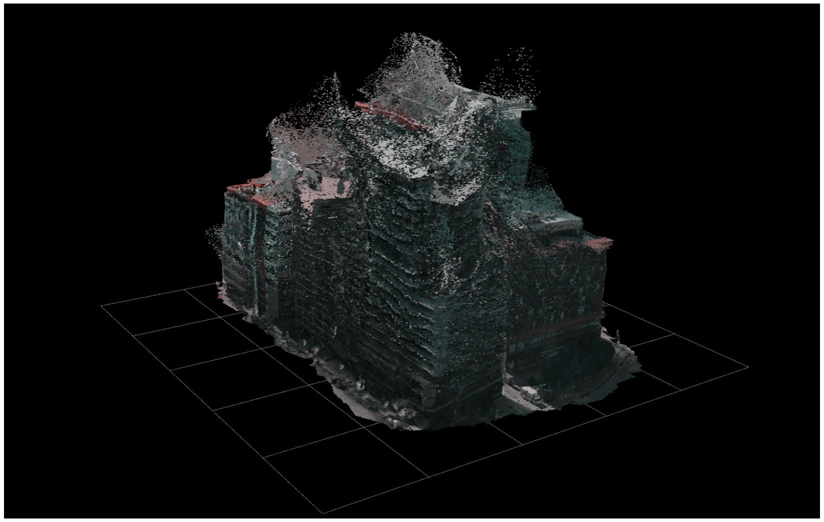
➔ Nano ID: Generates unique IDs for documents and audio files uploaded to the server.

MACHINE LEARNING and AudioStellar

As mentioned above, we used AudioStellar to achieve the goal of recreating urban sounds that are fed on our own small dataset of field recordings, as well as data collected from the web application. AudioStellar’s clustering algorithm allowed us to have different clusters of sound extracted from the audio files. Also, AudioStellar’s built-in tools such as sequencers and effects allowed us to have in-software sound design and also randomized and mutated sequences over time to keep the sound fresh and new for each iteration.

MACHINE LEARNING in MAX/MSP

A Max/MSP patch was designed and developed for the receiving audio files from the device. Firstly, a Node server was running to receive the Argon device’s audio files and store them in the host computer. In addition to that, a Python script was continuously running that keeps track of the incoming files and stores them in a queue. The queue processing is (wait line = audio_file_duration + 15secs. which is the machine learning part’s duration). This script would trigger the Max/MSP patch with the correct audio file’s name to be played. The sound would then first be played in a normal manner and then slowly reverb + delay effects would take over to disintegrate and decay the original audio file, and ultimately it would be replaced with the Machine Learning model’s decoded sound. The model that we used was Isis trained by IRCAM and then processed through RAVE as an encoder/decoder. This model was useful because it was trained on speech and we could use it to replicate the speech-audio files that we'd receive from the device. This patch also sent OSC messages to TouchDesigner where the visuals were rendered for when the ML model was activated with its sound frequency sent over to TD for special visual effects.



VISUALS for INSTALLATION

In our initial midterm report, the visuals were engaging but deviated significantly from our original concept. They failed to align with our overall theme and lacked recognizability. Through trial and error, we refined our approach by utilizing a Blender add-on called Blosm. This tool allowed us to render buildings from Google Maps tiles into editable 3D meshes, which were then streamed into CloudCompare—a software for rendering and editing point clouds. Using CloudCompare, we calculated point clouds from the mesh and integrated them into TouchDesigner, where we created the overall displacement effects. Additionally, we implemented audio reactivity, incorporating input from both the device in use and the general soundscape. To further enhance the visuals, we expanded our scope by including a much larger tile of downtown Montreal. Over time, the visuals warped and meshed together, achieving a greater level of cohesion. The physical buildings of the visuals were meant to reflect where the largest clusters of sound were, due to still being in the earlier stages of sound collections placeholders were swapped in.

This refined process resulted in visuals that better aligned with the audio, successfully creating reusable and recognizable elements. The final outcome was both cohesive and thematically consistent with the overall project.

SOUND DESIGN/COMPOSITION

The vision for the composition was to create a dynamic soundscape that evolves autonomously. The goal was to design a system capable of generating an immersive auditory experience that enhances the installation without overshadowing other audio components. However, the first iteration required significant manual input to control parameters and trigger changes.

First Prototype

1. Sine tones with timbral change:

Multiple sine waves added together with timbral changes over time formed the base layer of the soundscape. The combination of harmonic and inharmonic frequencies created mostly pleasant, clear tones with evolving harmonic textures.

2. Reese Bass / “wobble” effect:

Using this method outlined by Tom Hall[1], a reese bass was added to provide a deep, resonant, and rhythmic low-frequency sound that would complement the higher sine tones. I manually adjusted the detuning amount and filter cutoff to create movement in the bass line.

3. FM Synthesis

A simple synth with frequency modulation, and an amplitude modulator. This was incorporated to enrich the overall soundscape with harmonic complexity and interest that stands out from the droning sine tones. The modulation ratio produces either harmonic inharmonic sidebands. I manually adjusted these parameters.

Final Iteration

The final version of the composition was an autonomous generative system. After clicking a few toggles and designating speed parameters, the system requires no intervention, resulting in a continuous and unpredictable auditory experience. It combined randomness and with a few deterministic rules like how often randomness occurred, to produce a dynamic, evolving soundscape. Key changes included:

1. Sine tones with timbral change:

Randomized panning for each of the eight oscillators was incorporated to give listeners a sense of movement and unpredictability depending on where they are in the installation space. This also resulted in more dramatic harmonic / timbral changes.

2. Reese Bass / “wobble” effect:

Automatic randomization of parameters (such as modulation depth, detuning, frequency, and panning) was added to maintain an evolving bass texture and ensure that the bass does not remain static. An OSC message was integrated to synchronize frequency changes with visuals in TouchDesigner

3. FM Synthesis

The carrier frequency of the FM synth is now driven by an automated stream of MIDI notes, adapted from Ben Johnson's "One-Button Piano" patch.[2] This output was then merged with the frequency signal of the reese bass, which acted as a modulating signal of the FM synth's frequency.

The modulation ratio and amplitude were also randomized to introduce variability. Control parameters like a minimum and maximum range and rate of change were added.

4. Pink and white noise

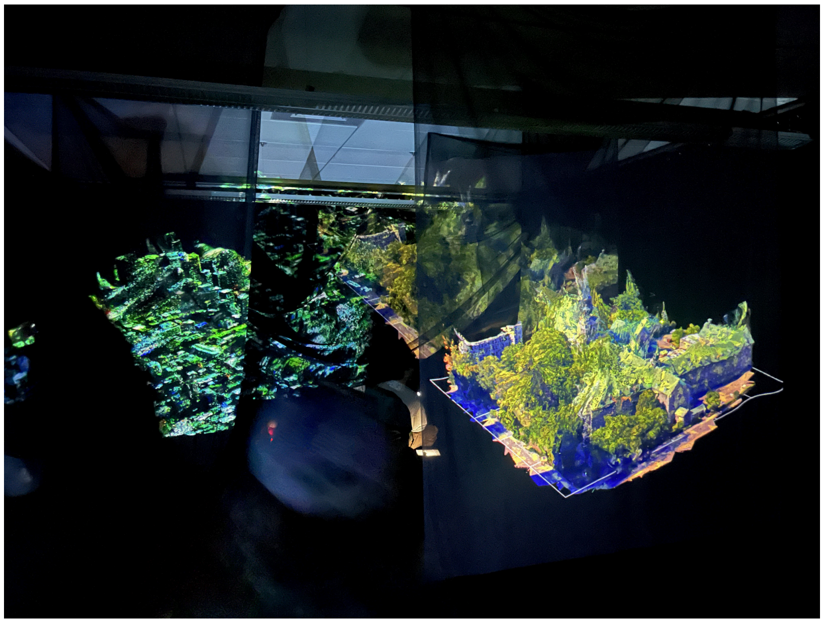
I incorporated pink and white noise to add texture and variation to the soundscape. A pink noise and white noise signal are each processed through a chain that normalizes, clips, and scales the signal. The scale is automatically randomized varying the relative loudness of each signal, with parameters to control the rate of change. The pink noise has a lower frequency and a smoother, more natural sound. The white noise contains all frequencies in the audible spectrum at equal intensity, resulting in a bright, sharp, and often "hissy" sound. The combined signals were automatically panned.

[1] https://www.tomhall.xyz/project/reese-bass-maxmsp/

[2] https://cycling74.com/articles/one-button-patching-challenge.

Final Artifact

Despite undergoing some changes during the prototype and development process, the core intention of Jamais-Vu remained consistent. The goal is to create a defamiliarizing audio-visual experience that prompts reflection on how technology and data shape our perception of the urban environment and its potential futures.



Observations

During the exhibition period, we noticed that first and foremost the audio queue needed more refinement as to how it's played out after people leave their messages. Our initial idea was to have this 35 second delay between each audio track, but we didn't keep in mind that the queue line might be a long one and people might not hear their audio tracks before leaving. This leaves some room for improvement and refinement as to how the queue is handled. An initial first solution could be to decrease the delay and merge audio files faster.

The device placement and general interaction seemed to be comprehensive enough for the participants and we'd like to reiterate on sound spatialization and dedicated spaces of the room being completely machine learning versus untouched areas. The idea of sitting down on the bench also worked out fine as some people took their time to sit down and take the place in and be fully immersed.

Also, the general movement around the space seemed to be successful and participants were encouraged by the layout to explore the space. Regarding the visuals, it was apparent that we still need some reiteration as to how we could make the audio-reactive aspect of it more apparent.

Ultimately, the web application's deployment was successful and we gathered a large amount of recording data to play with. The implied feeling of Jamais Vu was entirely present in the space and the immersed simulated urban life allowed some space to participants to reflect on the datafication and digitization of our everyday urban life. The very notion of our past, present and future regarding the human noise in urban spaces allowed an area of reflection for us the creators and the participants alike to pay attention to what noises are being made, and how these noises are mirrored digitally.

Future Directions

EXPANSION AND DIVERSIFICATION OF AUDIO COLLECTION

Expanding the audio sample collection by getting more parts of the city that we didn't manage to for the final artefact is important to realizing the full potential of this project. We aim to incentivize crowd contributions more and more to actively archive the urban soundscape.

BUILDINGS BETTER REFLECTED BY AUDIO CLUSTERS

As our sound database grows, we'll have more data on where most of the recorded audio files come from and we can include more visualized buildings for the mappings.

FINDING A PROPER SPACE

A larger and more adaptable space would enhance the experience of the installation. The current iteration, designed for the Holodeck, helps us visualize what it could look like in a bigger space, which would allow for more design choices. A much larger space would also allow for physical structures or props that would help evoke the feeling of navigating through the city. Given that this is a completely site-specific installation, we aim to find such spaces to hold our exhibition.

Bibliography

Baudrillard, Jean. 1983. *Simulations*. New York City, N.Y., U.S.A. : Semiotext(e), Inc.

Crawford, Lawrence. "Viktor Shklovskij: Différance in Defamiliarization." Comparative Literature, vol. 36, no. 3, 1984, pp. 209–19. JSTOR, https://doi.org/10.2307/1770260.

Fisher, Mark. 2014. *Ghosts of My Life: Writings on Depression, Hauntology and Lost Futures*. Winchester: Zer0 Books.

Ge, Xiao, Chunchen Xu, Daigo Misaki, Hazel Rose Markus, and Jeanne L. Tsai. 2024. "How Culture Shapes What People Want From AI." In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3613904.3642660.

Team Members and Associated Responsibilities, Tasks and Inter-Team Division of Labour.

Installation Design:

- This task includes the physical design of the space itself, material exploration, and also how the audience is guided through the experience.
- Members involved: everyone

DEVICE DESIGN/DEVELOPMENT:

- The handheld devices will allow installation visitors to record their voice or sounds and directly contribute to the soundscape in real-time.
- Members involved: Rebecca - Azhar

APP DESIGN/DEVELOPMENT:

- The app is the main online checkpoint where the participants can submit their recordings to the online database. This task included the design of the web checkpoint, development of the backend server and the database alike.
- Members involved: Louis

MACHINE LEARNING and Audio Synthesis:

- RAVE and Audiostellar will be used as the neural network system (and fine-tuned) to defamiliarize the recordings and collections. This task includes fine-tuning the model and developing the algorithms for the installation.
- Members involved: Kamyar

VISUALS for INSTALLATION:

- The visuals present at the installation will be mainly done by TouchDesigner. The projections will be mapped on sheet fabrics to create the illusion of physical structures. This task includes the initial development, as well as mapping to the installation space.
- Members involved: Noah

SOUND DESIGN/COMPOSITION and Synthesis:

- Supporting sound to accompany the installation to hold the general aesthetic together and guide the sonic experience, without overtaking the other sound elements.
- Members involved: Lydia