Critical Climate Machine: A Visual and Musical Exploration of Climate Misinformation through Machine Learning

GAËTAN ROBILLARD, ESIEE Paris - Université Gustave Eiffel, France and Arcanes - Université de Montréal, Canada
JÉRÔME NIKA, Ircam - Sorbonne Université, CNRS UMR STMS 9912, France

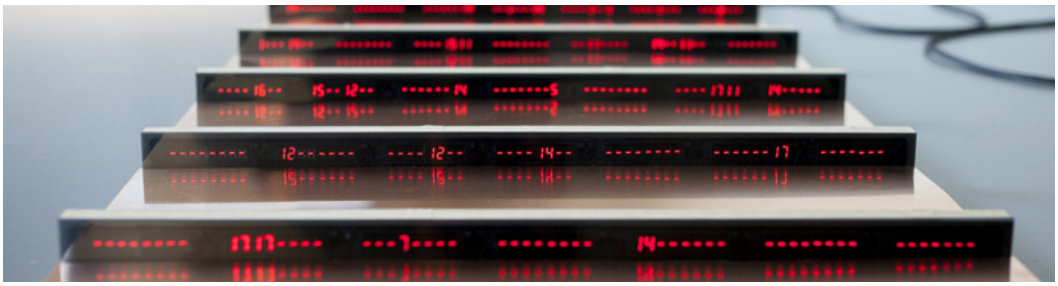

Fig. 1. Critical Climate Machine, detail view. The sculpture analyzes messages collected on social network X (Twitter). © Gaëtan Robillard.

Critical Climate Machine is a cutting-edge media art installation that critically exposes and quantifies mechanisms of climate change misinformation. Utilizing computational aesthetics across data, imagery, and sound, this work processes real-time data from X (Twitter) through a natural language processing learning model derived from cognitive sciences. It not only renders the statistical aspects of this data visually but also manifests its thermal effects. A unique audio dimension is introduced through dialogues between climate skeptics and climate advocates, processed by the generative machine learning (ML) algorithm Dicy2. These elements collectively shape the installation, each unveiling its distinctive algorithmic aesthetics and technical underpinnings. This paper concentrates on the dual application of ML algorithms: one for dissecting extensive online misinformation streams, and the other for creating climate-related dialogues. This dual approach opens a discussion on the mediation of climate, at the convergence of computational and physical realms. Our aim is to critically examine the role of ML technologies in crafting aesthetic experiences that resonate within scientific discourse and public debate on climate issues.

CCS Concepts: • **General and reference** → *Empirical studies*; • **Hardware** → *Temperature monitoring*; • **Information systems** → **Data extraction and integration**; • **Human-centered computing** → **Empirical studies in visualization**.

Additional Key Words and Phrases: machine learning, aesthetics, climate change, misinformation, visualization, sculpture, music, sound, dialogues, social network, Twitter, X, classification, generative composition

Authors' Contact Information: Gaëtan Robillard, ESIEE Paris - Université Gustave Eiffel, Noisy-le-Grand, France and Arcanes - Université de Montréal, Montréal, Canada, gaetanrobillard.studio@gmail.com; Jérôme Nika, Ircam - Sorbonne Université, CNRS UMR STMS 9912, Paris, France, jerome.nika@ircam.fr.

ACM 2577-6193/2024/8-ART56
https://doi.org/10.1145/3664215

ACM Reference Format:
Gaëtan Robillard and Jérôme Nika. 2024. Critical Climate Machine: A Visual and Musical Exploration of Climate Misinformation through Machine Learning. *Proc. ACM Comput. Graph. Interact. Tech.* 7, 4, Article 56 (August 2024), 11 pages. https://doi.org/10.1145/3664215

1 INTRODUCTION

Critical Climate Machine (CCM) is a media art project that reveals and quantifies mechanisms of misinformation about climate change[1]. The work employs computational aesthetics, in its data, image, and sound form. To this end, two machine learning (ML) algorithms, which process both climate-skeptic discourse and its science-based rebuttals were utilized. In this article, we will particularly highlight these.

Our research investigates the presentation of climatosceptic discourse and its critical counterarguments in contemporary media, aiming to explore the role of ML in shaping aesthetic experiences linked to scientific and public climate debates. Although existing literature addresses the reception of fake news [Wagner and Boczkowski 2019], our focus diverges to examine the application of ML in the artistic and technological mediation of climate issues within the informational sphere.

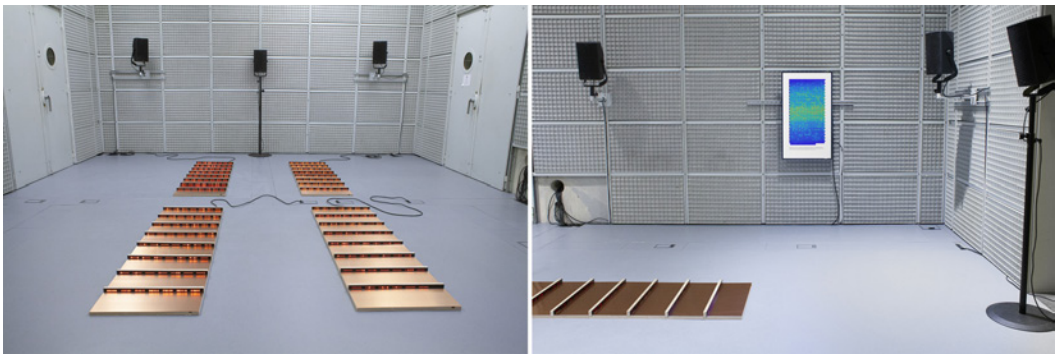

Fig. 2. Critical Climate Machine, installation view, IRCAM Centre Pompidou, Paris, 2022. © Gaëtan Robillard.

Firstly, we will present current research on both ML and algorithmic aesthetics in the field of visual arts and musical composition. Secondly, we will present CCM, including the adaptation of a learning model based on Natural Language Processing (NLP), which is trained to classify contrarian claims about climate change. We will detail how this model is used to process new data collected from the social network X (Twitter). Thirdly, we will detail the sound work, which is based on a participatory process. It employs recent developments of the generative ML algorithm Dicy2 in the field of musical interaction.

2 PRIOR RESEARCH

Today, ML algorithms are not only a major topic in computer science, but they are also intricately woven into the fabric of contemporary artistic exploration. In some cases, they are used to evaluate or generate aesthetic objects in a global economy marked by the generalization of data analysis [Arielli 2021]. In other cases, algorithms are examined for their agency in the creative process itself

[1]CCM is part of the MediaFutures project, funded by EU and Horizon 2020. It is the winner of the « BCS Futures Award » from the Lumen Prize, it also has been nominated by the STARTS Prize 2023. The work has been exhibited in venues such as : Université Gustave Eiffel (2024), IRCAM Centre Pompidou (2022), ZKM Center for art and media (2021-2022), Deutsches Museum (2021). Video report: https://vimeo.com/667971904.

[Boden and Edmonds 2019]. CCM resides at the intersection of data, ML, and generativity, drawing upon prior works that intricately intertwine creativity with these technological facets.

2.1 Waves in the Matrix

CCM hardware design derives from *The Wave in the Matrix*, an installation and a research by Gaëtan Robillard [Robillard et al. 2020] that examined recent storm waves in the North Atlantic, while stressing the digital means by which the signs of climate can be seen and perceived .

The installation implemented a model derived from ocean science that generates waves based on archived data and mathematical-statistical modeling. Coded in Python language, the model was distributed in a network of 32 processors embedded in sculpture modules (Figure 3), which together displayed the dynamics of a wave field in storm condition.

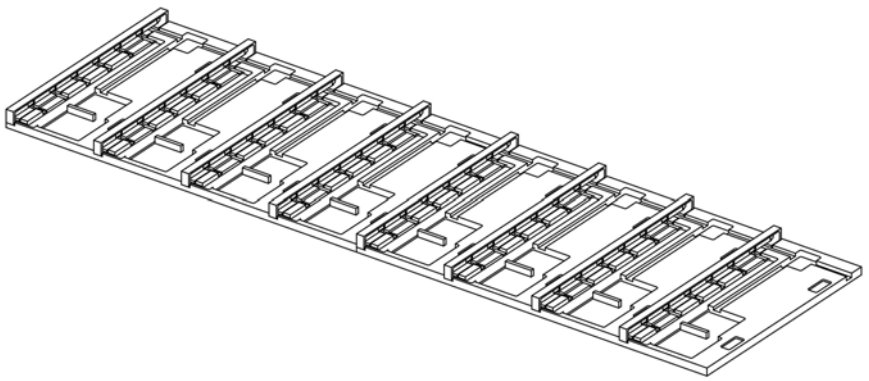

Fig. 3. Axonometric view of two modules, 200 x 60 x 4.5 cm, 2019. © Gaëtan Robillard.

The hardware configuration of the installation was based on a network of processors covered with copper and a matrix of 1,792 LED digits. Depending on the exhibition contexts, the temperature of the hardware varied, slightly altering the processor's speed. To amplify this, temperature of each processor was then regularly displayed by the sculpture, reflecting on algorithmic aesthtetics based on ecological dimensions.

The sonic dimension of the artwork consisted of an electro-acoustic performance by Aude Rabillon, based on voice. As a result, the artwork combined scientific visualization and a musical approach to climate change. CCM directly extends on this design and aesthetics.

2.2 Interactive music creation from a "memory" model

The purpose of the sound part of CCM is to confront raw data with their (re-)mediation by ML algorithms, by looking for ways for them to be distinguished or conversly hybridized. The Dicy2 library[2] implementing this research contains a collection of generative agents and tools for smart composition and human-machine interaction combining ML and generative processes with reactive listening modules to propose free, planned as well as reactive approaches to corpus-based generation [Nika et al. 2022].

The controls on the generative processes offered by Dicy2 make it possible to easily create large numbers of variations around the same explicit or underlying structures, introducing then novel

[2]Available at https://forum.ircam.fr/projects/detail/dicy2.

practices of "meta-composition" that will be discussed later. In the next section, we will focus on the new art installation CCM using two different types of ML models, including Dicy2 agents.

3 REVEALING MISINFORMATION ABOUT CLIMATE CHANGE

CCM is an installation made with a data sculpture based on ML algorithms, supervised or markovian, a visualization and a sound installation. The installation notably aims to explore the exposure of the social network X to climate skeptic discourse. A software[3] was implemented to render the results of the tweets analysis in the sculpture and a ludic card game was designed as a basis for workshops to collect audio material to be used in the sound installation.

3.1 Cognitive sciences

Exploration in cognitive sciences led to the discovery of a dataset intended to train ML models for classification of contrarian claims about climate change [Coan et al. 2021].

This training dataset originates from nearly 250,000 documents from prominent American conservative think tanks and contrarian blogs over 20 years. It consists of paragraphs labeled by climate-literate volunteers according to five known types of misleading arguments about climate change: "it's not happening", "it's not us", "it's not bad", "solutions won't work", and "climate science is unreliable"[4].

With this labeled dataset and by training ML algorithms, it becomes possible to explore large amounts of online data linked to climate skepticism. With CCM, we adapted the method from Travis G. Coan et al. to the analysis of X feeds. We designed a specific software to monitor tweets, and created a database to store the newly collected data alongside labels generated on the fly.

3.2 896 tweets

Our software collects tweets from 44 skeptical accounts linked to american conservatives lobbies. On top of the query, a list of 91 keywords are used to filter the claim research, with keywords such as "Antarctica", "IPCC" (Intergovernmental Panel on Climate Change) or "fossil fuel".

During inference, the sculpture calculates the probability of each input (tweet) belonging to the encoded categories (labels) identified in the training dataset (Figure 4). To do so, the algorithm first transforms each tweet into a quantity of numbers (a vector) and then uses a classifier (logistic regression).

The distributed data stream runs through each unit of the sculpture and updates until it stabilises. As a result the visual output of the installation becomes a landscape of "error codes", which are in fact labels indexes. Among the 896 tweets displayed by the sculpture simultaneously, the audience is invited to identify misleading tweets and map them to known contrarian arguments with the assistance of a code sheet (see Figure 5).

No text is displayed by the machine. However, by showing sequences of vector values and label indexes, the machine manifests the statistical nature of the ML algorithm operations. Note that the logistic regression algorithm cannot interpret text by its litterary meaning, but instead quantifies relationships between words and classes, according to word frequencies. Other limitations should be underlined. Beyond the potential for bias within the dataset, which warrants further investigation (e.g. tone, emojis, ponctuation), our model operates on a supervised learning framework. Consequently, its training may not fully capture the complete range of contrarian viewpoints on climate change.

[3]Available at https://github.com/robillardstudio/c-climate-public.
[4]The 5 top categories were subdivided into 27 subcategories.

Fig. 4. Critical Climate Machine, installation view at Deutshes Museum (ZukunftMuseum), Nürnberg, 2021. © Gaëtan Robillard.

-. none
1. ice isn't melting
2. we are heading into ice age
3. weather is cold
4. there is a hiatus in warming
5. sea level rise is exaggerated
6. extremes aren't increasing
7. it's natural cycles
8. there is no evidence for greenhouse effect
9. sensitivity is low
10. there is no species impact
11. co2 is not a pollutant
12. policies are harmful
13. policies are ineffective
14. clean energy won't work
15. we need energy
16. science is unreliable
17. climate movement is unreliable

Fig. 5. Critical Climate Machine, contrarian arguments labels.

Nonetheless, by exposing its algorithmic dynamic, and making its interpretative framework tangible, the sculpture takes on the value of a signal, highlighting ethical considerations in the

production of new data. But this is far from being the sole outcome. During the process, words not only become digits, but also turn into heat.

3.3 Patterns of Heat

Displayed on a vertical screen, and based on a grid of colors, a visualization witnesses the thermal evolution of the sculpture parts (Figure 6). It relies both on the temperature variation of processors embbeded in the sculpture, and the amount of detected misinformation at a given time. The warmer the installation and the higher the number of misinformation, the warmer the color tones. Through visualization, the ecological and social implications of ML techniques are highlighted. Color gradients signal the physical impact of computational processes, and eventually the way misinformation, but also big data, affects our social lives in unexptected ways.

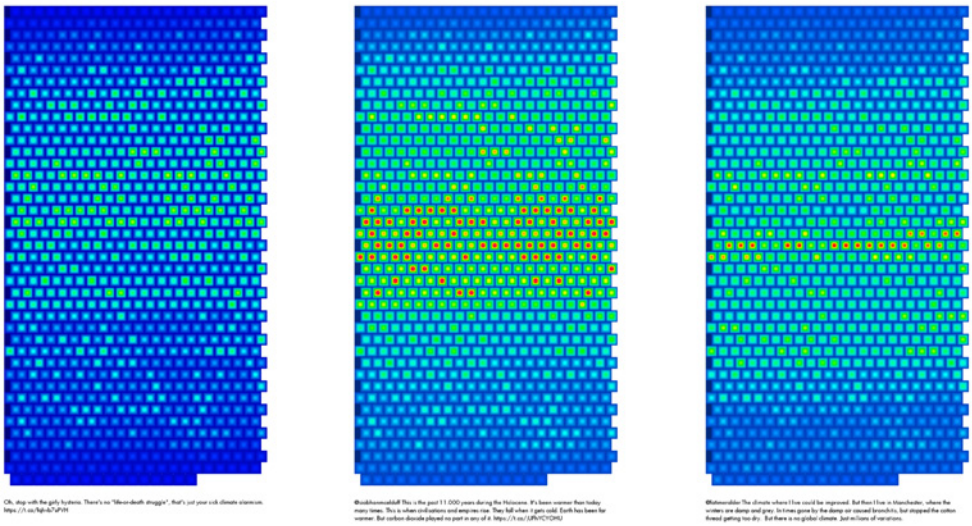

Fig. 6. “Patterns of Heat”, visualizations captured at three different stage. Each point represents one of the analysed tweets. Below each color grid, a randomly selected tweet is displayed. © Gaëtan Robillard.

To achieve this, experiments were conducted using a thermal camera aimed at the installation’s processors. Observations were made on the temperature variation during the inference process of collected tweets. We also compared the thermal effects of both the logistic regression classifier and the deep learning RoBERTa architecture [Coan et al. 2021], including electrical consumption. It is worth noting that the deep learning technique was significantly more resource-intensive. Finally, to enhance the visual display, we decided to represent the data in the form of a dot matrix, using a macro color gradient to visualize thermal variations, and micro gradients to visualize ML outputs.

4 VOICE, SOUND AND GENERATIVE MACHINE LEARNING

The sound installation is based on voice and dialogues, opposing climate skeptics and climate advocates. The setup includes a set of eight loudspeakers placed around the sculpture. It involves a critical and a formal work on speech. On the one hand deceptive claims are refuted, with the

help of scientific knowledge, and on the other hand authentic speech is echoed with generative speech. The nature of climate semiotics emerges both through contradictory debates and through compositional operations, conducted using generative ML techniques. The voices constituting the dialogues were written and recorded during several workshops based on the *Refutation Game*.

4.1 The Refutation Game

For younger generations, addressing climate change and combating misinformation is of utmost importance, and CCM is committed to tackling these challenges through research and fostering diversity. Central to this effort is the *Refutation Game*[5], a specially designed card game that facilitates structured debates by systematically refuting skeptical arguments.

This game plays a vital role in the creation and dissemination of the CCM project. It serves as an initial step in creating the soundscape for the installation and operates as a mediation tool during the exhibition phase, forming a cohesive circle within the process. The cards include a curated set of skeptical arguments from the training dataset mentionned in the above section, supplemented by refutations from Skeptical Science[6], a reputable science education organization. A final set of 40 short dialogues were written for the installation. From this initial set, 34 dialogues were then selected and subdivided to be performed by 6 voices (Figure 8) from students who already volunteered in the workshops.

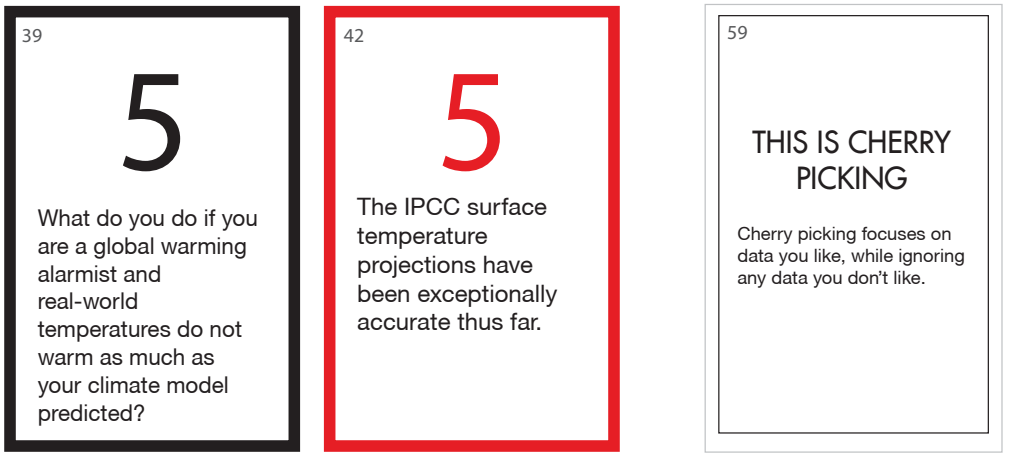

Fig. 7. The Refutation Game facilitates countering misleading claims. The figure depicts three different types of cards: a misleading argument, a refutation, and a reference in misinformation technique. © Gaëtan Robillard, Laurine Capdeville.

Overall, sixteen pupils and students participated in the production. Participants came from the international boarding high school of Paris (LIEP, public school) or from IMAC engineer program at Gustave Eiffel University. Located in the eastern suburbs of Paris, both institution welcome students from a variety of backgrounds and origins. Participation also extended to these students being directly involved in exhibitions at prominent venues like the IRCAM Centre Pompidou.

[5]Game design: Gaëtan Robillard and Laurine Capdeville
[6]See https://skepticalscience.com.

Fig. 8. Two students – Laëtitia Ngaha and Anastasiya Balan – recording dialogs between climate skeptics and climate advocates. © Gaëtan Robillard.

4.2 Sound work

Based on generative ML, the sound installation weaves the original dialogues with new ones; they are edited together and spatialized[7]. As a result, the audience experiences a movement between ordered meaning and unstructured language, revealing together ML processes and the nature of discourse on climate: its lexicon, its technicality, its repetitions. The overall composition traces the livelihood of a group of young people, echoed by syntehtic technics.

First, we introduced a permutation principle to explore contrasts in speech, like false arguments versus refutations, and authentic versus synthetic voices. The Dicy2 algorithm was then used to process these voices purely as musical elements (i.e. without semantic analysis). Our design ensured that the extent of modification by the generative ML process directly influences the spatial dynamics of the voices; the more they are altered, the more expansively they occupy the octophonic space (Figure 9). Finally, a player was designed to create dynamic sequences directly from a list of 102 generated files, representing approximately 1 hour and 20 minutes of material.

4.3 Generative Machine Learning with Dicy2

The Dicy2 generative processes are built from a "memory", a sound database serving simultaneously as learning ground for the models and as an acoustic reservoir for the synthesis. The challenge lies in the implementation of creative navigations through this memory to generate a new discourse. What we call "memory" is a model learned from an analysis of the data in a tailored automaton structure providing a graph connecting repeated sub-patterns in its sequence of labels [Nika et al. 2017].

[7]Sound design: Tony Houziaux and Dionysios Papanikolaou.

In CCM, the memory model is built from a representation of the recorded dialogues using audio descriptors [Peeters et al. 2011], that is to say time series calculated from the signal to characterize the temporal evolution of its acoustic and timbral characteristics.

This model provides a map of the memory's structure (repeated motives and variations regarding these audio descriptors) that can be walked through following different strategies. The memory can be browsed within a *free* generation run that will generate an audio sequence that respects the memory's hidden internal logic with no other specification than to be novel.

Alternatively, audio generation can be *guided* by a *scenario* querying the model, and that the output should match. This symbolic sequence is defined on the same label alphabet as the data representation used to learn the memory. In CCM, these labels represent therefore ranges of audio descriptors. The final audio output is a concatenation of the optimal sequence of audio memory chunks whose labels match the scenario.

Fig. 9. One of the 102 resulting octophonic files, represented as a waveform. © Gaëtan Robillard, Jérôme Nika, Tony Houziaux, Dionysios Papanikolaou.

The meta-composition paradigm changes radically: the user no longer explicitly defines a temporal evolution, but provides a corpus of sequences that will serve as "inspirations" to articulate the narrative of the generated outputs [Nika and Bresson 2021]. In CCM, the disruption of speech by generative processes was thus achieved by *free* runs in the memory model, recreating a voice without meaning but retaining a prosody. Other results have been obtained by chaining generative agents to hybridize structures and textures of different subsets of recordings, using for example the acoustic analysis of a climatoseptic argument as a *scenario* to follow for the generation of new material using its refutation as memory. For example, the engine in Figure 10 generates multiple audio files whose underlying scenario is generated by free generation in a selected sub-corpus of "structure" voices, and guiding the actual audio generation using a selected sub-corpus of "texture" voices.

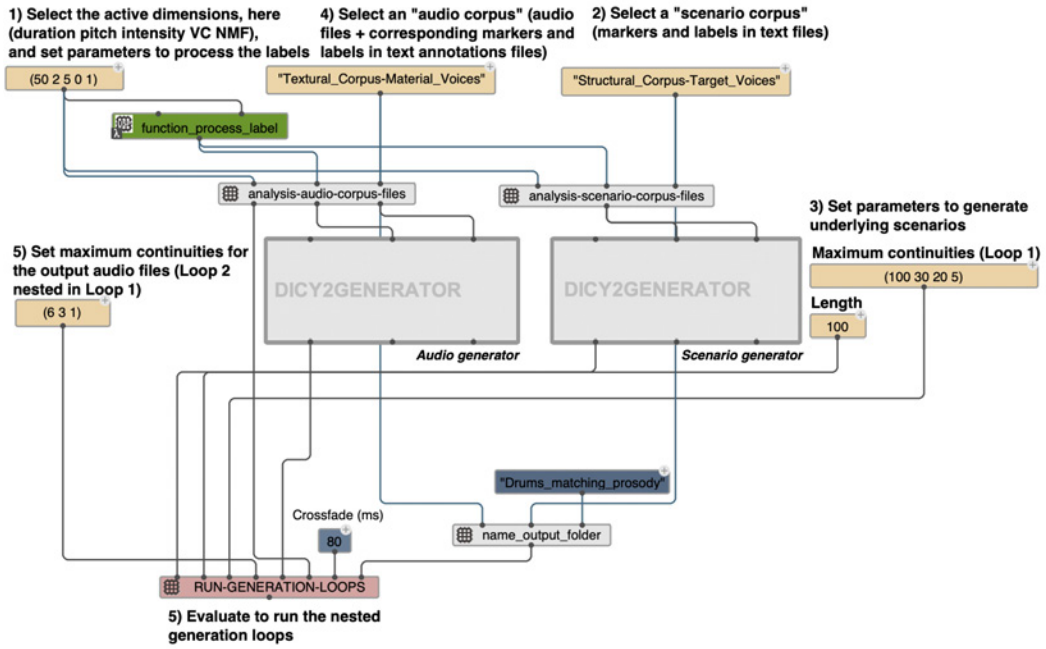

Fig. 10. Chaining two generative agents respectively in charge of the structure and the texture of the outputs. © Jérôme Nika.

5 CONCLUSION

In this project, we aimed to uncover the mechanisms behind climate misinformation by combining ML algorithms. In terms of functionnality, the installation identifies and categorizes new instances of climate-skeptic discourse emerging from X. CCM transforms this constant stream of data into a tangible visualization. Aggregating climate misinformation into a database strongly suggests that "AI techniques as art tools warrant further investigation regarding their effectiveness in countering misinformation within society" [Walker et al. 2023], offering a perspective on combating its spread.

We must also highlight that a number of public institutes and community centers have incorporated the *Refutation Game* into their workshops, demonstrating that our project meets genuine expectations in terms of climate and media literacy. Regarding future work, in addition to publishing the data collected through CCM[8], we plan to formalize an evaluation of the perception of climate misinformation, especially in the context of the *Refutation Game* (e.g. with likert scale).

Moreover, this research highlights the relation between NLP and generative musical research, especially when they share the same initial data. While computational logic in cognitive sciences aids in discerning truth, ML-based musical research ventures into exploring norms around climate mediation. Echoing Bernard Stiegler's philosophy of disautomation [Stiegler 2015], generative aesthetics facilitated by ML serves here as a counterpoint to technical determinism, promoting sustainable epistemologies. This implies significant perspectives in ML aesthetics, with a greater focus on both data and interactivity, and less on the automation of creativity itself.

The generative processes utilised for creating the sound in the installation belong to a current branch of generative musical systems research. Through the composition of high-level abstract

[8]Comprising 93,232 tweets from 2020 to 2023.

specifications and behaviors the research aims to develop new creative tools and practices. Instead of designing music by autonomous generative agents – found for example in deep learning techniques – we focus on practices where the compositional gesture remains fundamental. In CCM, concepts such as "memory" and "navigation" are valued as artistic leverage for experiencing timely issues in novel ways, encouraging both formalization and reflexivity throughout the process.

Consequently, we plan to further examine the interactions between ML aesthetics and climate discourse. This involves for instance devising a new visual and musical performance that utilizes real-time processes, incorporating live inputs from the audience when debating climate. Further along, our research could also benefit from venturing into new areas of misinformation and exploring new fact-checking methods (e.g., Google Fact Check Tools API). We can conclude that visual and musical paradigms have a common stake in the field of research in ML aesthetics. Considered together they allow progress, particularly in the critical and environmental dimensions.

REFERENCES

Emanuele Arielli. 2021. "Even an AI could do that". In *Artificial Aesthetics: A critical guide to AI, media and design*. Manovich and Arielli, 26.

Margaret A. Boden and Ernest A. Edmonds. 2019. *From Fingers to Digits: An Artificial Aesthetic*. MIT Press, Cambridge, MA.

Travis G. Coan, Constantine Boussalis, John Cook, and Mirjam O. Nanko. 2021. Computer-assisted classification of contrarian claims about climate change. *Scientific Reports* 11, 1 (Nov. 2021), 22320. https://doi.org/10.1038/s41598-021-01714-4

Jérôme Nika and Jean Bresson. 2021. Composing Structured Music Generation Processes with Creative Agents. In *2nd Joint Conference on AI Music Creativity (AIMC 2021)*. 12.

Jérôme Nika, Ken Déguernel, Axel Chemla, Emmanuel Vincent, and Gérard Assayag. 2017. Dyci2 agents: merging the "free", "reactive", and "scenario-based" music generation paradigms. In *International Computer Music Conference*.

Jérôme Nika, Augustin Muller, Joakim Borg, Gérard Assayag, and Matthew Ostrowski. 2022. *Dicy2 for Max*. Technical Report. Ircam UMR STMS 9912. https://hal.science/hal-03892611

Geoffroy Peeters, Bruno L. Giordano, Patrick Susini, Nicolas Misdariis, and Stephen McAdams. 2011. The Timbre Toolbox: Extracting audio descriptors from musical signals. *The Journal of the Acoustical Society of America* 130, 5 (11 2011), 2902–2916. https://doi.org/10.1121/1.3642604 arXiv:https://pubs.aip.org/asa/jasa/article-pdf/130/5/2902/15297399/2902_1_online.pdf

Gaëtan Robillard, Caterina Calgaro, and Emmanuel Creusé. 2020. *La vague dans la matrice*. Rapport de recherche 2020:02. Université de Lille, Lille. https://hal.archives-ouvertes.fr/hal-03133380

Bernard Stiegler. 2015. *La Société automatique. 1. L'avenir du travail*. Fayard, Paris.

María Celeste Wagner and Pablo J. Boczkowski. 2019. The Reception of Fake News: The Interpretations and Practices That Shape the Consumption of Perceived Misinformation. *Digital Journalism* 7, 7 (Aug. 2019), 870–885. https://doi.org/10.1080/21670811.2019.1653208 Publisher: Routledge.

Johanna Walker, Gefion Thuermer, Julian Vicens, and Elena Simperl. 2023. AI Art and Misinformation: Approaches and Strategies for Media Literacy and Fact Checking. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society (AIES '23)*. Association for Computing Machinery, New York, NY, USA, 26–37. https://doi.org/10.1145/3600211.3604715