INTERPRETATIVE DATA SONIFICATION: USING LLMS TO INTERPRET DATA AND GENERATE CONTINUOUS SOUNDSCAPES AT THE SYDNEY OPERA HOUSE

Rodolfo OCAMPO[1], **Oliver BOWN**[1], **Brendan WRIGHT**[2,1], **Justin SHAVE**[2], and **Caroline PEGRAM**[2]

[1]**University of New South Wales**, Sydney, Australia
[2]**Uncanny Valley Music**, Sydney, Australia

ABSTRACT

This paper describes a generative music work, 'Music of the Sails' developed as a commission for the 50th anniversary of the Sydney Opera House. The artwork consisted of a generative composition that responded to the building's data continuously in real-time for one month. The GPT-4 large language model (LLM) was used as an interpretative interface between the live data and a generative music engine. The data included information from the building's climate and energy systems, such as temperature and energy use, and its roster of concerts. The composition consisted of a generative arrangement of readymade music assets, including field recordings, AI-generated music, and human-composed music. Our approach, called Interpretative Data Sonification, leverages LLMs to interpret the building's data and use it to drive a generative soundscape. Music of the Sails was live-streamed on the Sydney Opera House website continuously for 30 days (744 hours) and exhibited as an audiovisual installation over three days. This research explores the roles that generative artificial intelligence (AI) can play in the interpretation of data as sound. We describe challenges faced around LLM output fixation and primacy bias and how future work can address these.

1. INTRODUCTION

This paper describes a generative music work, Music of the Sails, commissioned by the Sydney Opera House for its 50th anniversary. The brief was to use AI to bring the building to life, to reveal it not just as a place of performance but as a performer itself. In particular, they wanted to showcase the systems that inhabit the building, especially its custom seawater cooling system. To achieve this, we had access to near real-time data from systems that regulate the temperature of the building, air quality, water, energy and the monthly schedule of performances across the six venues inside the Opera House.

Building on previous work [1] we developed an approach combining data sonification, soundscape design and generative AI. Data sonification transforms data into sound to convey information sonically for various purposes, including data exploration, entertainment, and artistic expression [2, 3]. It often involves mapping numerical or categorical values to musical qualities such as tempo, pitch, or key [4]. However, these approaches often face the challenge of meaningfully conveying the qualities of the data sonically [5].

In contrast, human musicians transpose complex phenomena into musical compositions with layers of meaning. Classical approaches to the musical representation of the environment are interpretive, not direct. A classic example is Vivaldi's Four Seasons, capturing nuanced qualities of the changing seasons through music [6]. We sought to develop an approach that both reacted to data but that could integrate elements of human compositional skill and creative interpretation to richly convey the underlying phenomena behind the data. We used the capability of LLMs to manipulate simple concepts to interpret the data and then drive a music generation engine designed and curated by a human composer.

In a previous installation [1], we developed a first approximation of this approach that generated soundscapes by matching data descriptions generated by LLMs to a set of human-created audio descriptor tags using word embeddings and cosine similarity. The current implementation evolves that approach. Instead of performing latent space matchings of word embeddings of data and audio labels, we leverage language models to not only generate interpretations of data, but also to make selections about how the soundscape generation engine will work. In this way, we engage the LLM's capabilities more deeply and employ it as a tool with greater potential creative agency. The LLM was primed with information about its role and the artistic goals of the installation, as well as context about the data, the Sydney Opera House, and the music generation engine, providing the basis for it to make decisions about what music elements to select.

Candy [7] suggests AI can take multiple creative roles, including tools, mediums, mediators and partners. As a tool, the system is used by a human for a specific purpose and its effectiveness relies upon human skill. As a mediator, the system enables complex relationships between people, technology and the environment. As a partner, the system acts with a degree of agency, influencing the creative output non-deterministically.

In our installation, the AI seemingly expresses elements of each of these roles. It served as a tool to generate a human readable interpretation of data, as a mediator between the complex systems inhabited by and inhabiting in the building, and as a partner that made selections about the creative output experienced by the audience. However, in each of the roles, we encountered limitations, particularly in the tendency to fixate around certain outputs and the exhibition of primacy bias. In this paper, we discuss these limitations and possible avenues to address them.

2. RELATED WORK

The direct creation of soundscapes or music from data has been explored in a variety of contexts, from scientific exploration to artistic expression. Artists such as Marty Quinn's [8] and Andrea Polli [9] have explored sonification of data from various scientific fields, including ice core, radar, DNA, seismic, and solar wind data, with both aesthetic and data analysis goals.

More recently, several approaches have explored Natural Language Processing approaches to perform mappings between language and music. For example, Krol et al. [10] explored the use of GPT-3 to generate musical explanations, showcasing the potential of LLMs in creative musical contexts. Successful text-to-music generation has rapidly emerged in recent years. Companies like Mubert [1] combine user prompts, latent embeddings, and human-composed sound clips to generate music. This approach shares similarities with our work in terms of using semantic matching between text and sound.

In this context, Pigrem et al. [11] proposed that data sonificiation and soundscaping are two sonic art forms that have developed partly in isolation but that share important similarities, and thus proposed the concept of datascaping, where data sonification serves as a narrative device in soundscape composition. Their approach uses recorded sounds augmented with real-time sonified elements. Our approach is similar to datascaping, where our goal is to let data elements drive a rich, human-composed sonic experience. However, our approach also integrates the use of LLMs as a bridge between the two.

In our previous work [1], we developed Semantic Sonification, which involves a more direct mapping between data descriptions and selected audio elements. In the current approach, which we term Interpretative Data Sonification, we replace cosine similarity matchings of multimodal word embedding tags with an LLM that can drive the soundscape generation process based on an interpretation of the data, the soundscape engine and our artistic goals. We also involve a human compositional process more deeply, creating many fixed elements for the soundscape, along with AI-generated music loops, and field recordings inside the Opera House building.

3. IMPLEMENTATION

3.1 The Opera House Systems

The Opera House is renowned for its sophisticated self-regulation and energy efficiency systems [12]. It can be understood as an adaptive system that seeks to maintain homeostasis across multiple variables, including temperature, air quality, energy use, and water use. To achieve this, it employs a complex set of interconnected monitors and controls.

For example, the building uses seawater from the harbour in its cooling system. The water is circulated to cool the building and then returned to the harbour at a temperature that does not disrupt the ecosystem. This cooling system is powered by a set of six chillers that are activated in stages based on the cooling needs of the building. Like other buildings it also regulates CO2 levels to maintain acceptable air quality. It draws in outside air to reduce CO2 exhaled by people.

The Opera House contains several venues, including the Concert Hall, the Utzon Room, the Drama Theatre, the Playhouse, and the Studio. When these spaces fill up, such as during performances, the CO2 levels and temperature can rise quickly. In anticipation of heating, the building preemptively cools the venues before performances begin. To compensate for increased CO2 the building's systems draw in more outside air, which makes cooling less efficient than using recycled air. Consequently, the building is always in a state of dynamic adaptation, responding to both internal and external conditions.

3.2 The installation

Our installation needed to reflect the dynamic, continuous nature of the building's adaptive systems. It connected to these systems in real-time, generating a continuous soundscape that played 24 hours a day for 30 days over a livestream and as part of a 3-day audiovisual installation at the TIDE room inside the Sydney Opera House.

The building moves through different states throughout the day, and our goal was to generate music that reflected each of these states. For example, when the building was hot and busy, we wanted the soundscape to be lively, full of sounds, and to convey a sense of warmth. Conversely, during quiet periods in the middle of the night with no performances, we wanted the soundscape to feel calm and cool.

3.3 Interpretative Data Sonification

We received the data from the building each hour in the form of a CSV file containing minute-by-minute data points. These data were preprocessed to a simplified summary, formatted as a string and passed to the LLM in a prompt along with context about the meaning of variable names and ranges.

We then asked the LLM to generate an interpretation of the data that succinctly described what it said about the state of the building. We then gave the LLM information about the Memu music generation engine (see below), and asked it to select parameters controlling its state.

3.3.1 The Memu Music Generation Engine

Memu is a music generation system that combines ready-made stems with rule-based manipulation. As part of our installation, we utilized a variety of sound sources, including custom compositions by the fourth author, recordings

[1] www.mubert.com

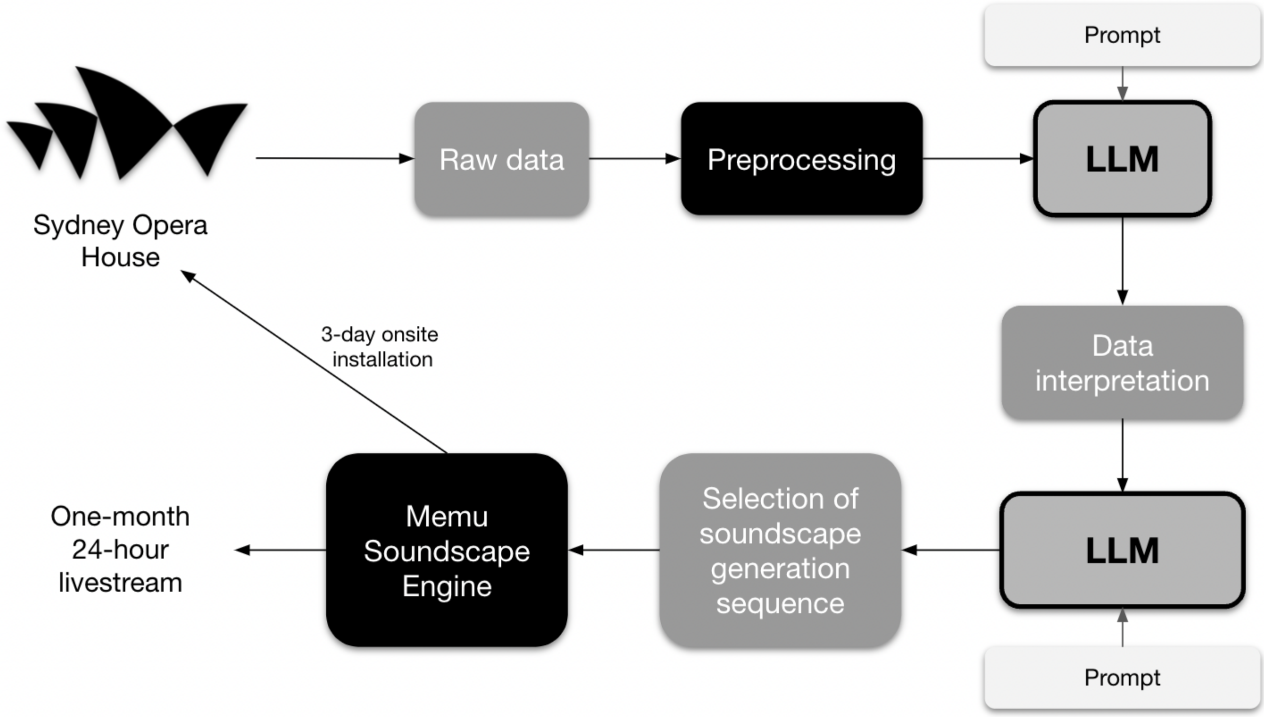

Figure 1. Flow of information for the installation, from data ingestion to soundscape generation.

Figure 2. The recordings from the Sydney Opera House building included sounds from the working infrastructure, which became part of the generative soundscape

from the Sydney Opera House's Grand Organ, field recordings from around the building, such as the chillers, water pumps, and electrical systems, as well as variations of these sounds generated by AI processes.

3.3.2 Sequences

Memu generates music in different modes of operation, termed *sequences*. Each sequence involves a set of combinatorial rules that seek to evoke a particular emotion, producing never-repeating 5-minute tracks that get interwoven to produce a continuous soundscape.

Each sequence was extensively labeled by the composer-engineer with the feeling it aimed to convey. Table 1 shows six examples out of the 22 crafted sequences.

The LLM was asked to interpret the data from the building and then select three sequences that would match the building's state, from a list of all possible sequences. This was done once per hour for 744 hours.

3.3.3 Data Collection and Preprocessing

The data from the Sydney Opera House consisted of minute-by-minute raw data from each room's regulating systems, including CO2 levels, room temperature, water temperature, ongoing events, as well as energy consumption from the building's chillers.

Table 2 shows an example of part of the data.

These data were formatted as JSON and passed hourly to the LLM as part of the prompt string.

3.3.4 Data Interpretation

We asked the LLM to generate a succinct description that captured the state of the building based on the incoming data. For example, for the data in table 2, the following description was generated.

Generated text:

> *Late evening at the House finds three chillers powering the cooling system, keeping the Drama Theatre and the Playhouse primed for imminent performances. The Joan Sutherland Theatre and other venues remain in a state of*

Table 1. Sequence labels. Sequences are series of rules that generate unique soundscapes. The labels are human-written descriptions of the sequences.

No.	Description
1	The music exudes feelings of joy and optimism. The harmonious melodic fragments invoke a sense of contentment and happiness, creating a serene and uplifting mood that resonates with feelings of hope and positivity.
2	The musical sequence evolves from serenity to vibrancy, with a shift in tonal dynamics and rhythmic intensity. It encapsulates the subtle transition from tranquility to lively anticipation.
3	In this sequence we reach an unwavering stability, offering a rhythmic continuity that balances sense of tranquil constancy and fervent activity.
11	Harmonious layers intertwine in a relentless, dynamic equilibrium, generating an atmosphere of thoughtful, energized tranquility.
21	The rhythmic experience unfolds in slow, disjointed cadences, each beat a spectral echo weaving through a dimension of languid resonance and ephemeral pulses.

tranquility, their temperatures and CO2 levels stable under a watchful system.

3.3.5 Soundscape Generation

Based on that description, we asked the model to select three sequences of operation for the Memu generative music system that best accommodated the building's state.

We also asked to justify how these soundscapes were suitable, and the justification was shown on screen as part of the visual elements of the installation.

Generated text:

Inspired by these data, the rendered soundscape encompasses the vibrancy of the theatre performances and the tranquility of the building's operation. The harmonious melodies reflect the joy of the shows, the transition to vibrancy mirrors the anticipation of the audience, while the constant rhythmic elements mirror the building's stable operation.

Selected soundscapes ids: 1, 2, 3

3.4 Audiovisual Installation and Live Stream

Visual elements were also included as part of the soundscape generation. The visuals were created using the generative video system Deforum, and involved five-minute clips of merging visuals that include motifs from the Sydney Opera House, themes from the event calendar, and different moods and themes corresponding to the music labels.

Table 2. Selection of Opera House data elements.

Data Point	Value
Active Chillers	Ch1, Ch2, Ch5
Chiller Stage	Stage3
Chiller Load (Ch1)	1112.883 kW
Chiller Load (Ch2)	964.2559 kW
Temp. (Concert Hall)	22.16°C
Temp. (Opera Theatre)	22.25°C
Water Temp. (Ch1)	15°C
Water Temp. (Ch2)	16°C
CO2 (Concert Hall)	451.4 ppm
CO2 level (Concert Hall)	Normal
CO2 (Opera Theatre)	406.8 ppm
Valve (Concert Hall Cold Water)	Stable
Valve (Opera Theatre Hot Water)	Stable
Temp. Oscillation (Concert Hall)	Stable
Event (Playhouse)	Twelfth Night

Overlaid on these visuals, the outputs from the LLM were displayed at regular intervals. These include the data interpretation and a musical explanation of how the generated soundscape reflected the data qualities. Lastly, graphs of the data were also displayed at regular intervals.

Figure 3 shows stills from the live stream displaying the generated text interpreting the data, and the data history for different venues, which appeared on screen at regular intervals.

When text or data was not being displayed, a stream of generative video visuals was shown, as shown in Figure 4.

4. DISCUSSION

A total of 744 hours of live data from the Sydney Opera House were used to drive the generation of music, live-streamed online on the Sydney Opera House website, and displayed as part of an audiovisual installation at the TIDE Room for three days, inside the House. To hear the soundscape and see the data and interpretations generated by the LLM model, the reader can access a 24-hour archive and a highlights reel from the entire month [2] .

The work demonstrated how an LLM could perform interpretation between a live data stream and a generative musical work. The use of an LLM as an interpreter or mediator afforded a different approach to data soundscape creation which was not constrained to establishing direct mappings from data to sound parameters. Instead, it enabled us to work in a more traditional mode of creative interpretation, freely developing musical material that reflected our impressions of the building and its various states and activities, and then orchestrating the ways in which the LLM selected this content. The design of prompts, including information about the building data, the set of music possibilities to choose from, and the creative objectives, enabled us to control the context that informed how the LLM interpreted the data. However, in practice the work encountered the natural limitations of what cur-

[2] https://stream.sydneyoperahouse.com/music-of-the-sails

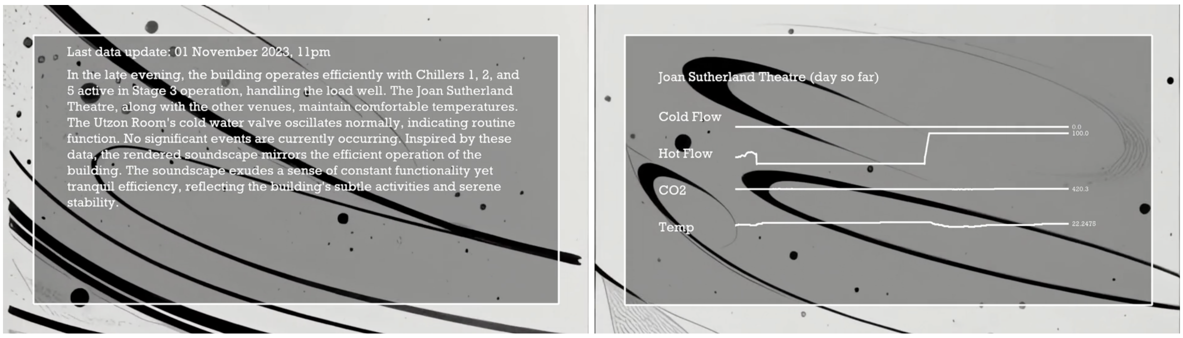

Figure 3. (Left) Text generated by the LLM interpreting the data and explaining the generated soundscape. (Right) Data visualization displayed during the installation, showing real-time data from the Sydney Opera House.

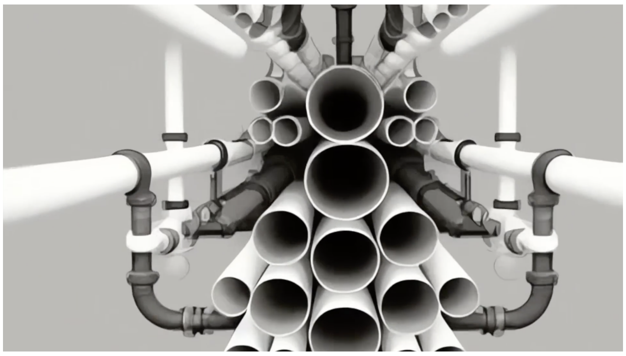

Figure 4. A still from the generative audiovisual stream

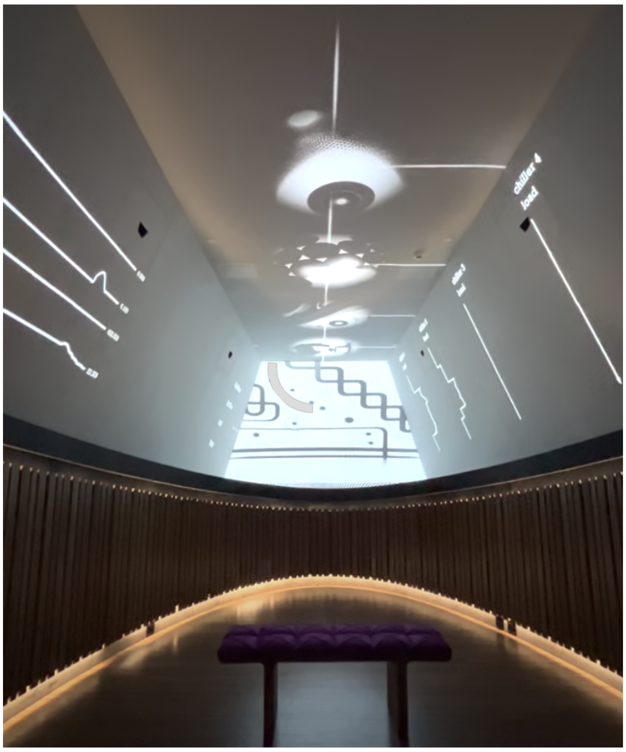

Figure 5. Photo of the onsite installation inside the TIDE room at the Sydney Opera House

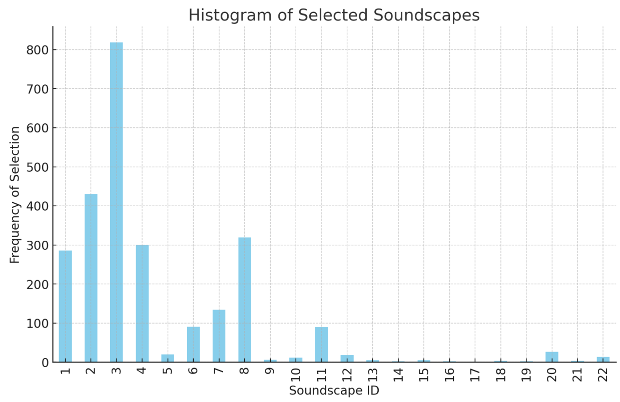

Figure 6. Histogram displaying sequence selection frequency

rent LLMs can do given this kind of task.

What was particularly exciting about this approach was the idea that, in principle, one could then maintain an ongoing dialogue with the LLM as it processed the data. Thus over the 744 hour performance it could be possible to give feedback and additional creative input. This was not practically possible in the present work due to processor constraints.

4.1 Limitations and Future Research

We identified limitations in our current approach. One of the objectives of our installation was to produce variety in soundscape generation. However, after examining the record of the selected soundscape sequences by the LLM, we found that it tended to fixate on certain sequences.

Figure 6 shows a histogram of the selected sequences for the entire month. As we can observe, the model tended to focus on a narrow number of sequences.

This can be due to several reasons.

4.1.1 Lack of Data Variety

The Sydney Opera House is a well-regulated building that maintains a steady equilibrium. Therefore, the model tended to select soundscape sequences that were tagged with words that reflected this stability, as well as positive terms that reflected excitement for events or coming down from activity. In contrast, it did not select sequences for soundscapes tagged with words that reflected instability, or negatively-valenced terms.

Table 3 shows the tags for the most frequently selected sequences.

Table 3. The most frequently selected sequences for soundscape generation

No.	Description
1	The music exudes feelings of joy and optimism. The harmonious melodic fragments invoke a sense of contentment and happiness, creating a serene and uplifting mood that resonates with feelings of hope and positivity.
2	The musical sequence evolves from serenity to vibrancy, with a shift in tonal dynamics and rhythmic intensity. It encapsulates the subtle transition from tranquility to lively anticipation.
3	In this sequence we reach an unwavering stability, offering a rhythmic continuity that balances sense of tranquil constancy and fervent activity.
4	This sequence gradually transitions into serenity as its rhythmic complexity diminishes, leaving behind a lingering sense of calmness and grace.
8	This music provides a soothing backdrop, while descending melodies evoke a sense of tranquil descent. This musical passage carries an air of introspection and peaceful reflection, offering a harmonious journey back to a state of quietude and contentment.

Table 4 shows the tags for the sequences that were not selected, which include more words referring to tension, disquiet and melancholy.

Figure 7 shows a wordcloud with size weighted by frequency of all the descriptions generated by the LLM in response to the Opera House data. As we can see, these words align more closely with the descriptions shown in table 3, corresponding to the most frequently selected sequences.

4.1.2 Primacy Bias

However, another factor that might have influenced the fixation was a well-documented bias in large language models: primacy effect or primacy bias. Several authors have reported this bias on large language models, where they tend to select the first elements in a list [13, 14]. This primacy effect is similar to primacy biases exhibited by humans and it is currently an active area of research since many of the applications of LLM's involve selecting elements from lists in semantic-based tasks. Primacy bias may have significantly influenced how the LLM selected options from our list of music possibilities.

Table 4. Sequences that were not selected at all by the system

No.	Description
14	This sequence, while ethereal and luminous, emits waves of unsettling and disconcerting harmonies, casting a celestial, yet uneasy resonance through the escalating rhythmic landscape.
15	The music pulses in a dynamic equilibrium, its harmonies interlaced with intricate, relentless patterns, creating an otherworldly tapestry that is both luminous and subtly disquieting, creating a dance balance between comfort and unease.
16	The rhythms start to unwind, soothing the previous tension and disquiet, allowing the ethereal harmonies to envelop the soundscape in a calming, albeit still subtly unsettling, embrace.
17	This sequence piece unfolds like a haunting, introspective whisper, weaving through the shadows with its melancholic harmonies and enveloping atmospheres.
18	This sequence commences a swell into reflective majesty, intensifying melancholic echoes and enveloping the senses in a crescendo of solemn magnificence.
19	The harmonies resonate within a majestic equilibrium, each somber note hovering in a monumental rhythmic dance of enduring resonance, creating feelings of solemn grandeur.
21	The rhythmic experience unfolds in slow, disjointed cadences, each beat a spectral echo weaving through a dimension of languid resonance and ephemeral pulses.

4.1.3 Dialogic Interaction to Address these Limitations

We believe there are two main ways to address the fixation on sequence selection in our Interpretative Data Sonification pipeline. Feedbacking the history of its previous selections (history awareness) and engaging in a dialogue in real-time to steer the LLM's behaviour.

4.1.4 History Awareness:

Feedbacking the history of its previous selections to the model, along with a specific prompted goal of avoiding repetitiveness, we believe we could partly address the issue of fixation and primacy effects. In this installation, we did not include such previous history to save inference cost and time. However, with increasing context token windows and the ability to include moving windows that progressively summarise inference history, including large amounts of previous context becomes more feasible.

4.1.5 Dialogic Interaction:

On the other hand, being able to engage in a dialogue with the LLM and say things like: "you have selected sequence three too many times, try selecting others" could more eas-

Figure 7. Word cloud of the descriptions generated by the LLM in response to the Opera House data, with size weighted by frequency.

ily help address the fixation. During the installation, we partly modified the prompts in real-time to try to steer the model in a particular direction after observing undesired behaviour. A step beyond this would include a chat interface where the human creator would engage in a running dialogue with the LLM performing the Interpretative Data Sonification matching and therefore be steered more effectively without prompt trial and error. This, combined with history awareness, both of previous selections and conversations with the user, would provide better adherence to diversity goals.

Bown et al. [15] explore the concept of dialogic interaction as a mechanism to enable more effective human-AI co-creative interactions. We believe future installations would benefit from more explicitly enabling dialogic forms of collaboration with generative AI models.

Such a dialogue would elevate the AI model even more to a co-creative and active role, in which it is able to adjust to changing objectives and even be flexible with regard to the initially established creative objectives.

5. CONCLUSIONS

In this paper, we presented 'Music of the Sails', a generative soundscape installation developed for the 50th anniversary of the Sydney Opera House that combines data sonification, human soundscape design, and generative AI to create a sonic celebration of the House's changing states. Our approach, termed Interpretative Data Sonification, leverages large language models (LLMs) to interpret the building's real-time data and drive a generative soundscape engine. The project explored the roles of LLMs as creative tools in sonic installation, in which human compositional skill serves as the backbone of a generative data-driven process.

The experimental work revealed limitations of this method, particularly the LLM's tendency to fixate on certain musical sequences, possibly due to primacy bias and limited data variability, which constrain the variety of soundscapes. To address these issues, we propose feedbacking the history of the model's previous selections to encourage diversity and enable real-time dialogic interaction for dynamic adjustments to the model's behavior. Despite the outlined challenges, 'Music of the Sails' contributes to the growing body of research on AI-driven creative systems, highlights the potential of LLMs as realtime interpreters or mediators between complex data and artistic expression, and opens up new modes of interaction between human artistry, generative AI, data and sonic art.

6. ACKNOWLEDGMENTS

We would like to thank the Sydney Opera House team: Stuart Buchanan, Melinda Dransfield, Rob Millet, and James Archer.

We also thank the entire Music of the Sails team: Chloe McFadden (generative visuals), Charlton Hill (management), Stephen Krol (AI-music stems) and Felix Wallis (additional music).

We acknowledge the Gadigal people, the traditional custodians of the land where the Sydney Opera House stands.

7. REFERENCES

[1] R. Ocampo, J. Andres, A. Schmidt, C. Pegram, J. Shave, C. Hill, B. Wright, and O. Bown, "Using gpt-3 to achieve semantically relevant data sonificiation for an art installation," in *Artificial Intelligence in Music, Sound, Art and Design: 12th International Conference, EvoMUSART 2023, Held as Part of EvoStar 2023, Brno, Czech Republic, April 12–14, 2023, Proceedings*. Springer, 2023, pp. 212–227.

[2] D. Rocchesso, S. Serafin, F. Behrendt, N. Bernardini, R. Bresin, G. Eckel, K. Franinović, T. Hermann, S. Pauletto, P. Susini *et al.*, "Sonic interaction design: sound, information and experience," in *CHI'08 extended abstracts on Human factors in computing systems*, 2008, pp. 3969–3972.

[3] T. Hermann, A. Hunt, and J. G. Neuhoff, *The sonification handbook*. Logos Publishing House Berlin, 2011.

[4] F. Grond and J. Berger, "Parameter mapping sonification," in *The Sonification Handbook*. Logos Publishing House Berlin, 2011, ch. 15, pp. 363–397.

[5] S. Barrass, "A perceptual framework for the auditory display of scientific data," *ACM Transactions on Applied Perception (TAP)*, vol. 2, no. 4, pp. 389–402, 2005.

[6] M. Talbot, *The Vivaldi Compendium*. Boydell Press, 2011.

[7] L. Candy, "Creating with the digital: Tool, medium, mediator, partner," in *Interactivity, Game Creation, Design, Learning, and Innovation*. Springer International Publishing, 2020, pp. 13–28.

[8] M. Quinn, "The climate symphony and other sonifications of ice core, radar, DNA, seismic and solar wind data," in *Proceedings of the 2001 International Conference on Auditory Display, Helsinki Finnland*, 2001.

[9] A. Polli, "Atmospherics/Weatherworks: A Multi-Channel storm sonification project," *International Conference of Auditory Display*, Jul. 2004.

[10] S. J. Krol, M. T. Llano, and J. McCormack, "Towards the generation of musical explanations with gpt-3," in *Artificial Intelligence in Music, Sound, Art and Design: 11th International Conference, EvoMUSART 2022, Held as Part of EvoStar 2022, Madrid, Spain, April 20–22, 2022, Proceedings*. Springer, 2022, pp. 131–147.

[11] J. Pigrem and M. Barthet, "Datascaping: Data sonification as a narrative device in soundscape composition," in *Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences*. New York, NY, USA: ACM, Aug. 2017.

[12] Sydney Opera House. (2023) Sustainability. [Online]. Available: https://www.sydneyoperahouse.com/about-us/in-the-community/environmental-sustainability

[13] J. E. Eicher and R. Irgolič, "Compensatory biases under cognitive load: Reducing selection bias in large language models," *arXiv preprint arXiv:2402.01740*, 2024.

[14] Y. Wang, Y. Cai, M. Chen, Y. Liang, and B. Hooi, "Primacy effect of ChatGPT," in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, H. Bouamor, J. Pino, and K. Bali, Eds. Singapore: Association for Computational Linguistics, Dec. 2023, pp. 108–115.

[15] O. Bown, K. Grace, L. Bray, and D. Ventura, "A speculative exploration of the role of dialogue in Human-ComputerCo-creation," *ICCC*, 2020.