

# SoundSignature: What Type of Music Do You Like?

**Brandon James Carone**

Department of Psychology, New York University  
Music and Audio Research Laboratory (MARL), NYU  
Center for Language Music and Emotion (CLaME), NYU  
New York, New York  
[bcarone@nyu.edu](mailto:bcarone@nyu.edu)

**Pablo Ripollés**

Department of Psychology, New York University  
Music and Audio Research Laboratory (MARL), NYU  
Center for Language Music and Emotion (CLaME), NYU  
New York, New York  
[pripolles@nyu.edu](mailto:pripolles@nyu.edu)

**Abstract**—*SoundSignature* is a music application that integrates a custom OpenAI Assistant to analyze users' favorite songs. The system incorporates state-of-the-art Music Information Retrieval (MIR) Python packages to combine extracted acoustic/musical features with the assistant's extensive knowledge of the artists and bands. Capitalizing on this combined knowledge, *SoundSignature* leverages semantic audio and principles from the emerging Internet of Sounds (IoS) ecosystem, integrating MIR with AI to provide users with personalized insights into the acoustic properties of their music, akin to a musical preference personality report. Users can then interact with the chatbot to explore deeper inquiries about the acoustic analyses performed and how they relate to their musical taste. This interactivity transforms the application, acting not only as an informative resource about familiar and/or favorite songs, but also as an educational platform that enables users to deepen their understanding of musical features, music theory, acoustic properties commonly used in signal processing, and the artists behind the music. Beyond general usability, the application also incorporates several well-established open-source musician-specific tools, such as a chord recognition algorithm (CREMA), a source separation algorithm (DEMUCS), and an audio-to-MIDI converter (basic-pitch). These features allow users without coding skills to access advanced, open-source music processing algorithms simply by interacting with the chatbot (e.g., can you give me the stems of this song?). In this paper, we highlight the application's innovative features and educational potential, and present findings from a pilot user study that evaluates its efficacy and usability.

**Keywords**—*Music Information Retrieval (MIR), Natural Language Processing (NLP), Music Analysis Tools, Digital Signal Processing (DSP)*

## I. INTRODUCTION

The rise of digital media platforms in the 21st century has made music more accessible than ever, dramatically transforming how we consume and engage with it [1]. Globally, people listen to over 18 hours of music per week, facilitated by music streaming services and social media platforms that allow music lovers to explore, share, and enjoy millions of songs at their fingertips [2-4]. This shift from physical to digital media has not only profoundly affected the way we listen to music [1, 5, 6], but also the type of music we listen to: our daily listening habits are now heavily influenced by recommendation algorithms developed by streaming services [7-10]. While music

streaming platforms collect a diverse plethora of data to create compelling and individualized experiences, users are often unaware of the reasons behind the music suggestions made by a recommender system [11, 12]. To inform users about their musical behaviors, products such as Spotify Wrapped and Apple Music Replay, for example, focus on painting a picture of each user's listening history throughout the year. However, they fail to dive deeper into what encourages users to return to the songs and artists they love most and the reasons why individuals use different types of music in their daily lives [13, 14].

Over the last two decades, advancements in Music Information Retrieval (MIR) techniques have opened several new avenues for computer programmers and computational musicologists to decompose, analyze, and alter music and audio signals [15]. One of the problems with these existing tools is that they require a degree of technical expertise that can become an entry barrier for those who lack a background in signal processing and/or in coding skills. While remarkable efforts have been made to use music as a contextual tool to facilitate the learning of the technical foundations required to understand signal processing (e.g., capitalizing on music to teach linear algebra, Fourier analysis, etc.) [16], many times these interactive lessons still require coding skills. This barrier contributes to the significant gap in the user's understanding of the acoustic characteristics of the music they love, as many music listeners might be eager to explore the deeper aspects of their favorite songs but lack the skills to do so effectively. Moreover, in spite of recent efforts in promoting public musicology [17] and musical analysis, both traditional and computational musicology [18] remain confined to academic or professional settings, limiting accessibility for general users. Thus, there remains a critical need for resources that are both easy to use and computationally efficient, removing the technical barriers while at the same time providing advanced music-related insights.

This paper introduces *SoundSignature*, a music analysis web-application that leverages state-of-the-art MIR and Natural Language Processing (NLP) technologies. After a user uploads a sample of their favorite songs, the application extracts various musical and acoustic features from each musical excerpt and feeds the results to a customized OpenAI Assistant [19]. The chatbot then combines - in a way that is both accessible and user-friendly - the extracted features (e.g., BPM, Spectral Centroid, Spectral Flux, etc.) with its extensive knowledge of

the artists behind the songs to provide insights into the user's musical preferences. This system not only identifies musical elements but contextualizes them, enabling users to explore their musical tastes on a nuanced level, aligning with current research that emphasizes personalization and context-awareness in music recommendation systems [20]. *SoundSignature* focuses on providing users with a personalized overview of the specific acoustic and musical characteristics that make up their favorite songs, akin to a musical preference personality report. We believe this feature sets *SoundSignature* apart from other products that simply provide a summary of the listening preferences of the user (e.g., Spotify Wrapped). This emphasis on personalization aims to provide users with the feeling of being catered to and understood, thus enhancing their music experience and providing a description of their musical taste that goes beyond stating which musical genres or artists they like or have listened to the most. Another unique characteristic of *SoundSignature* is that it provides a brief description of the lyrical content and cultural context of each song and artist, respectively. Lyrics are often overlooked in music recommendation systems and personalized products like Spotify Wrapped. Therefore, incorporating OpenAI's vast knowledge of song lyrics and information about the artists might help users to better understand the meaning or emotion behind their favorite music [20, 21].

In addition to offering insights, *SoundSignature* encourages users to delve deeper into the analyses performed by allowing them to ask the chatbot further questions about the songs or the analyses described. Much like Mueller's approach to simplifying signal processing with musical coding examples [16], here we seek to further democratize signal processing by not only using music as a contextual tool to catch the users' attention and motivate them, but also by employing the users' favorite songs to enhance understanding of complex acoustic and/or musical concepts like spectrograms, spectral flux, or pulse clarity, among others.

Finally, we implement various state-of-the-art open-source tools such as music source separation and chord recognition algorithms, empowering users without coding skills to engage directly with advanced music processing algorithms (e.g., a user can ask the app to extract the stems of a song). Cloud computing, with its ability to offload processing and provide real-time, scalable access to complex systems, offers a promising solution to make sophisticated music analysis accessible to a broader audience. This aligns with the ecosystem provided by the *Internet of Sounds* (IoS) and the *Internet of Musical Things* (IoMusT) frameworks, where interconnected devices and cloud-based platforms enable seamless access to complex music processing and analysis tools [22-25]. The MIR analyses used by *SoundSignature* are performed via a *Streamlit* web application, which facilitates real-time interaction and cloud-based processing. Utilizing cloud infrastructure ensures scalability and accessibility, allowing users to interact with the system from any device with an internet connection [22, 26, 27]. By simplifying access to these tools, *SoundSignature* not only contributes to the growing IoS ecosystem, but also aims to facilitate creative exploration and experimentation for users with higher musical knowledge but limited coding skills.

This project is motivated by three aims: 1) to provide music enthusiasts, regardless of their musical knowledge or technical skills, with deep, personalized insights into their musical tastes; 2) to serve as an educational platform for users to learn about complex musical and acoustic concepts; and 3) to be a one-stop shop for musicians to use musical tools such as music source separation and chord recognition algorithms that are otherwise inaccessible to users with limited coding skills. To evaluate the efficacy and usability of *SoundSignature* we conducted a pilot user study. The methodology of this study, alongside detailed descriptions of the technology and algorithms employed, are explored in the following sections. This approach substantiates the application's capabilities and sets the stage for discussing its broader implications in reshaping how we interact with and understand music.

## II. SYSTEM ARCHITECTURE

The proposed system is composed of three main components: 1) a set of feature extractors, 2) tools for manipulating and processing the songs, and 3) an AI assistant that ties the tools together, synthesizing the extracted features with its knowledge of music to provide insights for the user. The entire system is presented to the user through a cloud-based web application. Thus, *SoundSignature*, situated within the broader context of the IoS and the IoMusT, interacts with musical data in a way that aligns with the IoS vision of connected, interactive, and intelligent sound devices. The following subsections describe the feature extraction pipeline, the additional tools integrated for musicians, the role of the AI assistant, and the web interface used to deliver the system.

### A. Musical / Acoustic Feature Extraction Pipeline

Feature extraction was performed with a variety of different music and signal processing packages, such as *librosa* 0.10.1 [28] and *essentia* 2.1.b6.dev117 [29] running in Python version 3.9. We began by using *librosa* to extract the audio signal and sampling rate, and then ran the short time Fourier transform on the audio signal. Following that, we output the spectrogram for the user to see. While many acoustic and/or musical features can be calculated, we capitalized on previous MIR, music psychology, and music neuroscience literature to select a few features representing core elements of music [30-32]. Specifically, we focused on features clustered across four categories: Tempo and Rhythmicity, Harmony and Melody, Timbre and Texture, and Energy and Dynamics. For Tempo and Rhythmic features, we focused on beats per minute (BPM; the speed or pace of the music) and pulse clarity (clarity of the rhythmic pulse or beat). For Harmonic and Melodic features, we extracted the key and mode (e.g., C major or A minor), as well as the key strength (clarity and stability of the detected musical key). For the Timbre and Texture features, we extracted the spectral centroid ("center of mass" of the spectrum, indicating brightness), spectral bandwidth (the width of significant frequency bands, indicating sound fullness), and spectral flux (the rate of change in the power spectrum, indicating musical onsets or texture changes). Finally, for the Energy and Dynamics features, we extracted RMS energy (average power of the audio signal independent of human perception) and loudness (psychoacoustic model that incorporates frequency weighting to reflect the human ear's varying sensitivity to different

frequencies) [29]. The system’s reliance on acoustic and musical feature extraction aligns with the interdisciplinary domain of semantic audio [25, 33], which uses signal processing and machine learning techniques to extract structured, meaningful information from music. By applying these techniques, *SoundSignature* provides users with high-level interpretations of their music preferences, facilitating human interaction with audio data through accessible, interpretable insights. This is aligned with the increasing demand for semantic audio features in the context of the IoS [22].

### B. Tools for Musicians

In addition to the features extracted from each song, we have also included in the app some state-of-the-art music tools that can be especially useful for musicians and other individuals with a high musical knowledge (e.g., composers). We include tools for chord identification and stem and midi extraction. Specifically, the app includes *Meta’s Hybrid Transformers for Music Source Separation (HT-DEMUCS)* [34], the *Convolutional and Recurrent Estimators for Music Analysis (CREMA)* chord recognition algorithm [35], and *Spotify’s Basic-Pitch audio-to-midi converter* [36]. Users can call the functions from these packages from the prompt by using specific words such as "stems," "chords," or "midi" (e.g., "can you give me the stems of this song?"; "create a midi file from this song."; "what are the chords in this song?"). After mentioning one of these keywords, a dropdown menu appears, and you select which of the MP3s you are referring to. Chords are output in a chart with "Start Time", "End Time", "Chord", and "Confidence" as the column headers, whereas the stems and midi files are both playable and downloadable from the web application (See Supplementary Materials for example).

### C. OpenAI Assistant

In order to interpret the users’ requests and the output of the extracted acoustic and musical features, we created a custom OpenAI Assistant running on the gpt-4o model by utilizing the OpenAI API [19]. Whereas the ChatGPT prompt is rather general, allowing it to adapt to any query a user might input, our customized assistant was informed that it is an advanced music analysis tool. Here is a condensed version of the prompt that we used:

"You are a sophisticated music analysis tool designed to offer personalized insights into users' musical preferences and what their favorite songs reveal about them. Users upload their favorite songs (format: SongName\_ArtistName.mp3), and you analyze the similarities and differences using acoustic and musical features extracted with Python packages like *librosa* and *essentia*. Translate these technical features into relatable music characteristics, comparing and contrasting songs to highlight what defines the user's tastes. Emphasize patterns that indicate a preference for specific musical styles or elements and how these may reflect their personality. Provide insights into what users might seek in music, drawing from your knowledge of the artists and songs for a comprehensive analysis. Use varied descriptors to avoid repetition and adjust the complexity of the language based on user preferences, starting with simplified explanations. Conclude each analysis by asking if the user needs further clarification or more detailed technical explanations."

It is important to note that while all of the previously mentioned features are manually extracted from the MP3s in Python, the cultural and historical contexts and lyrical content are pulled from OpenAI's knowledge base, which was most recently updated in October 2023 [19].

### D. Building the Website

To create the *SoundSignature* website and avoid translating the Python code to html, we utilized *Streamlit* [37], a Python package meant to host data science and machine learning applications. Users can edit some model parameters on the web application, enter their own custom OpenAI API Key, speak instead of type, or upload images, among others. The homepage instructs users to upload their songs in the specified format (SongName\_ArtistName.mp3; see Figure 1A). After the user uploads their music, the analysis function is called, and the results are displayed for the user to view (see Figure 1B). Once the analyses have finished running, the extracted features are fed to the chatbot, whereby it then produces its response (see Figure 1CDE). When coding the website, we attempted to maintain a simple user interface while ensuring that things ran smoothly, much like the ChatGPT website. See also the Supplementary Materials for examples of full responses of the app. For those interested in using the *SoundSignature* application, please reach out via email to bcarone@nyu.edu to discuss access. We aim to accommodate users within the constraints of our current resources.

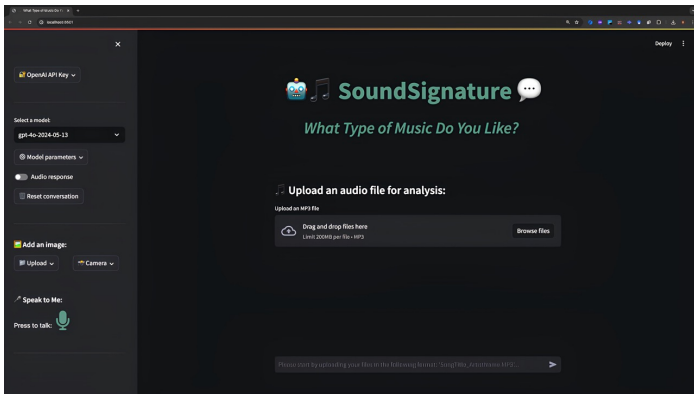
## III. PILOT USER STUDY

To evaluate the efficacy and usability of *SoundSignature* we conducted a pilot user study in which we collected quantitative and qualitative responses from participants who used the app.

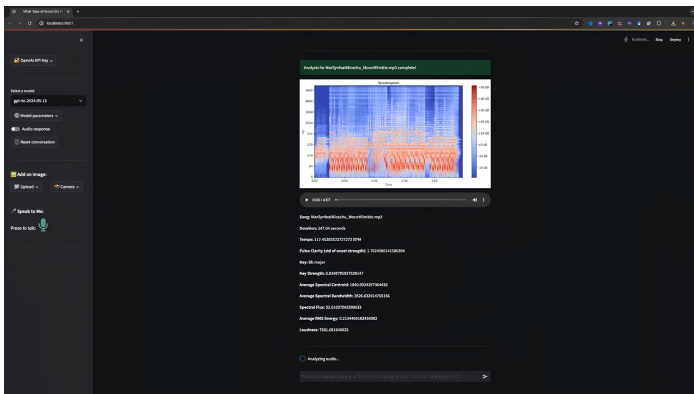
### A. Participants

A total sample of 20 participants (7 females, mean age = 33.65, sd = 13.22) completed the study. For participant demographics, the sample consisted of 60% White, 25% Other, 5% East Asian, 5% Black or African American, and 5% who preferred not to answer. Regarding ethnicity, 70% were Hispanic or Latinx. In terms of education, 35% were enrolled in or had completed a graduate program, 45% held a Bachelor's degree, 10% had some college education, and 10% completed some of high school. Informed consent was collected from each participant, and the study was completed in accordance with the NYU Institutional Review Board. Participants were paid for their participation or completed the study on a volunteer basis. Participants completed a series of questionnaires aimed at characterizing their musical background. Specifically, participants completed the Barcelona Music Reward Questionnaire (BMRQ), which measures sensitivity to music reward. The BMRQ was used to ensure that no participants had specific musical anhedonia (BMRQ Score < 63) [38]. Note that people with musical anhedonia show decreased activity within the reward system in response to music, and this would likely affect the user's perception of the application [39]. None of the participants recruited had specific musical anhedonia, with the average sensitivity to musical reward of the participants (BMRQ mean plus sd:  $86.9 \pm 8.2$ ) being slightly higher than that of the general population which is usually around 80 [38, 40]. Participants also completed the Goldsmith Music Sophistication Index - Musical Training subscale (Gold-





1A. Homepage of SoundSignature



1B. SoundSignature: Analysis Pipeline following upload

Welcome to your personalized music analysis session! Today, we'll explore what your favorite tracks from Voxtro, Mount Kimble, King Harvest, The Black Eyed Peas, Noah Cyrus ft. Leon Bridges, Future Islands, and Death Cab for Cutie tell us about your musical tastes. From the soulful rhythms to the intricate melodies, let's dive into what makes these songs resonate with you.

### 1. Tempo and Rhythmic Elements

- TheStartofSomething\_Voxtro: 107.67 BPM, Pulse Clarity: 1.52
- MarilynfeatMicachu\_MountKimble: 117.45 BPM, Pulse Clarity: 1.76
- DancingintheMoonlight\_KingHarvest: 136 BPM, Pulse Clarity: 1.95
- Alive\_TheBlackEyedPeas: 117.45 BPM, Pulse Clarity: 1.99
- July\_NoahCyrusft.LeonBridges: 71.78 BPM, Pulse Clarity: 1.95
- Moonlight\_FutureIslands: 129.20 BPM, Pulse Clarity: 1.34
- IWillFollowYouIntoTheDark\_DeathCabForCutie: 78.30 BPM, Pulse Clarity: 1.39

Insights: Your song choices show a preference for a wide range of tempos, from the slow and contemplative "July" by Noah Cyrus ft. Leon Bridges to the more upbeat "Dancing in the Moonlight" by King Harvest. The pulse clarity varies, suggesting you enjoy both distinct rhythmic patterns and more complex, less predictable rhythms.

### 2. Harmonic and Melodic Elements

- TheStartofSomething\_Voxtro: Key: Ab major, Key Strength: 0.88
- MarilynfeatMicachu\_MountKimble: Key: Bb major, Key Strength: 0.83
- DancingintheMoonlight\_KingHarvest: Key: C minor, Key Strength: 0.84
- Alive\_TheBlackEyedPeas: Key: B minor, Key Strength: 0.82
- July\_NoahCyrusft.LeonBridges: Key: C major, Key Strength: 0.82
- Moonlight\_FutureIslands: Key: C major, Key Strength: 0.92
- IWillFollowYouIntoTheDark\_DeathCabForCutie: Key: F major, Key Strength: 0.88

Insights: You have a balanced taste for both major and minor keys, indicating a flexibility in enjoying both uplifting and melancholic melodies. The strong key strengths suggest a preference for songs with clear harmonic foundations.

### 3. Timbre and Texture

- TheStartofSomething\_Voxtro: Brightness: 1717.40, Fullness: 2064.50, Texture Changes: 82.55
- MarilynfeatMicachu\_MountKimble: Brightness: 1840.09, Fullness: 2526.63, Texture Changes: 52.61
- DancingintheMoonlight\_KingHarvest: Brightness: 1911.03, Fullness: 2113.46, Texture Changes: 38.41
- Alive\_TheBlackEyedPeas: Brightness: 1883.01, Fullness: 2374.61, Texture Changes: 95.93
- July\_NoahCyrusft.LeonBridges: Brightness: 1709.83, Fullness: 2105.62, Texture Changes: 53.67

Please start by uploading your files in the following format: "SongTitle\_ArtistName.MP3"

1C. Tempo and Rhythmic Elements, Harmonic and Melodic Elements, and Timbre and Texture

- Alive\_TheBlackEyedPeas: Brightness: 1883.01, Fullness: 2374.61, Texture Changes: 95.93
- July\_NoahCyrusft.LeonBridges: Brightness: 1709.83, Fullness: 2105.62, Texture Changes: 53.67
- Moonlight\_FutureIslands: Brightness: 1048.21, Fullness: 1328.36, Texture Changes: 65.21
- IWillFollowYouIntoTheDark\_DeathCabForCutie: Brightness: 1678.43, Fullness: 2089.26, Texture Changes: 32.68

Insights: The songs you enjoy feature a range of timbral qualities, from the bright and full textures of "Dancing in the Moonlight" to the more subdued and mellow "Moonlight" by Future Islands. This suggests an appreciation for diverse soundscapes and a willingness to explore different sonic environments.

#### 4. Energy and Dynamics

- TheStartofSomething\_Voxtro: RMS Energy: 0.22, Loudness: 7648.78
- MarilynfeatMicachu\_MountKimble: RMS Energy: 0.21, Loudness: 7581.68
- DancingintheMoonlight\_KingHarvest: RMS Energy: 0.11, Loudness: 2316.57
- Alive\_TheBlackEyedPeas: RMS Energy: 0.27, Loudness: 11550.54
- July\_NoahCyrusft.LeonBridges: RMS Energy: 0.18, Loudness: 4178.69
- Moonlight\_FutureIslands: RMS Energy: 0.18, Loudness: 5354.00
- IWillFollowYouIntoTheDark\_DeathCabForCutie: RMS Energy: 0.10, Loudness: 2591.28

Insights: Your preferences span a wide range of energy levels, from the high-energy "Alive" by The Black Eyed Peas to the more subdued "I Will Follow You Into The Dark" by Death Cab for Cutie. This indicates a versatility in your listening habits, appreciating both energetic and mellow tracks.

#### 5. Cultural and Historical Context

- Voxtro: Known for their indie pop sound, Voxtro's music often features catchy melodies and introspective lyrics.
- Mount Kimble: Pioneers in the post-dubstep genre, their music blends electronic and acoustic elements.
- King Harvest: Their classic hit "Dancing in the Moonlight" is a staple of 70s pop-rock.
- The Black Eyed Peas: Known for their energetic and genre-blending music, often incorporating elements of hip-hop, pop, and electronic.
- Noah Cyrus ft. Leon Bridges: This collaboration brings together contemporary pop with soulful influences.
- Future Islands: Known for their synthpop sound and emotive performances.
- Death Cab for Cutie: Their music often features introspective lyrics and a blend of indie rock and alternative influences.

Insights: Your selections reflect a broad appreciation for different musical eras and genres, from the nostalgic 70s pop-rock of King Harvest to the modern indie and electronic sounds of Mount Kimble and Future Islands.

Please start by uploading your files in the following format: "SongTitle\_ArtistName.MP3"

1D. Timbre and Texture (cont.), Energy and Dynamics, and Cultural and Historical Context

- Noah Cyrus ft. Leon Bridges: This collaboration brings together contemporary pop with soulful influences.
- Future Islands: Known for their synthpop sound and emotive performances.
- Death Cab for Cutie: Their music often features introspective lyrics and a blend of indie rock and alternative influences.

Insights: Your selections reflect a broad appreciation for different musical eras and genres, from the nostalgic 70s pop-rock of King Harvest to the modern indie and electronic sounds of Mount Kimble and Future Islands.

#### 6. Lyrical Analysis

- Voxtro: Often introspective and reflective, dealing with themes of love and personal growth.
- Mount Kimble: Lyrics can be abstract and experimental, fitting their innovative sound.
- King Harvest: "Dancing in the Moonlight" is a feel-good anthem about enjoying life.
- The Black Eyed Peas: Often feature themes of positivity, partying, and social issues.
- Noah Cyrus ft. Leon Bridges: "July" is a poignant reflection on relationships and personal struggles.
- Future Islands: Known for emotive and heartfelt lyrics, often dealing with love and longing.
- Death Cab for Cutie: Introspective and often melancholic, dealing with themes of love, loss, and existential reflection.

Insights: Your lyrical preferences lean towards introspection, emotional depth, and personal reflection, suggesting a tendency to connect with music on a deeper, more personal level.

#### 7. Conclusion (Overall Musical Preferences)

Based on our analysis today, your music preferences show a deep appreciation for indie and alternative genres with a tendency to enjoy both complex textures and clear harmonic foundations. Each song you've selected offers a unique story, mirroring aspects of your own musical journey. Your choices reflect a balance between energetic and mellow tracks, indicating a versatile and dynamic listening habit. The lyrical themes suggest a reflective and introspective personality, with a penchant for emotionally resonant and thought-provoking music.

#### 8. Disclaimer

It is possible for the interpretation of the analyses or for the analyses themselves to be incorrect. If you recognize any false information or if the chatbot has hallucinated in any way, please email [bcarone@nyu.edu](mailto:bcarone@nyu.edu) and let us know what happened!

#### 9. Asking for Further Input

Would you like any further information on the output, or would you like a deeper explanation of anything technical? I can use metaphors and real-life examples to explain more technical acoustic or musical features if needed.

Please start by uploading your files in the following format: "SongTitle\_ArtistName.MP3"

1E. Lyrical Analysis, Conclusion (Overall Musical Preferences), Disclaimer, Further Input

Figure 1. Screenshots from the SoundSignature web application during use

MSI-MT) [41], which quantifies the participants' amount of musical training received. The average musical training score was (mean plus sd)  $28.6 \pm 11.1$ , with this mean falling into the 53<sup>rd</sup> percentile of the general population. Participants' favorite music genres were diverse, although 55% placed rock, 35% placed jazz, and 25% placed pop and alternative among their top 3 favorite genres.

### B. Experimental Design

Participants were first presented with an online consent form. They were then provided with a short overview of the procedures. Once online written consent was obtained, participants provided demographic information and then completed the BMRQ and the Musical Training subscale of the Gold-MSI. Participants then provided us with the names of 5-10 of their favorite songs (number of Songs mean plus std:  $7.4 \pm 1.9$ ), which were obtained and formatted appropriately (i.e., SongName\_ArtistName.mp3). We then walked the participants through the web application interface and uploaded their selected songs. Finally, once the analysis was completed and the assistant had responded, the participants were told to read through the response and let us know if they had any questions. Participants were also encouraged to interact with the chatbot further, and were provided with sample questions to ask (e.g., "What do these song choices say about me as a person?", "Based on these song choices, which other songs would you recommend?", etc.). Once they were done asking questions, they were required to complete a Likert scale from 1 (Completely Disagree) to 7 (Completely Agree) to rate how much they agreed with the Conclusion section provided by the app regarding their overall musical preferences (see Figure 1E). Participants were also asked to provide a Likert scale (also between 1 and 7) rating how much they agreed with the following statement: "This application informed me of my music taste in a meaningful way." Finally, we asked participants for their general thoughts on the app, allowing them to write as much as they wanted in an open-ended way. In addition to these written responses, the experimenter also took notes of other comments the participants provided during their interactions with the app. The survey containing the demographic questions, BMRQ and Gold-MSI questionnaires, and the questions about *SoundSignature* were all created and presented with Qualtrics Version 06/2024 [42].

## IV. RESULTS

### A. Quantitative Responses

The user study conducted to evaluate *SoundSignature's* efficacy and usability illustrated that participants were significantly satisfied with the application. Specifically, the average Likert rating for the agreement with the Conclusion provided by the app regarding the user's musical preferences was  $6.3 \pm 0.9$ . In addition, participants also felt that the app informed them about their musical taste in a meaningful way, with the average agreement score for this statement being  $6.4 \pm 0.7$  (see Table 1).

### B. Qualitative Responses

Responses to the open ended question were mostly positive (see Supplementary Table 1 for all responses). For example, participants positive responses included: "...it is pretty accurate and provided an insightful view of my musical tastes that I have

not thought about before", "...felt very personal and told me a lot about myself and music preference", "...enjoyed reading the insights about myself based on my musical choices", "...I like how much more specific it gets once you ask follow up questions", among others. A theme that arose in the open-ended questions was the desire of users to know more about the relationship between the acoustic and musical characteristics of their favorite music and their mood and/or personality traits: "...what is the relationship between the kind of music that I listen to and my personality traits, mood, and feelings based on the tempo, spectral features, and lyrics...", "...could the app analyze the songs I have heard during a day and then tell me which kind of music and/or which kind of emotions I was looking for on that day?", "...what it could tell me in other areas of my life, such as movie taste, personality type...", "...tell me a bit more personal information, almost like a musical horoscope". One participant provided negative feedback, stating that "...most of the conclusions were correct and it informed me of my music taste but I think it doesn't work that well with Classical pieces...", while another participant mentioned that "...the first output is a bit generic".

TABLE 1

*Descriptive statistics describing how much participants agreed with the Conclusion section provided by the app regarding the user's overall musical preferences response, and how much they agreed with the statement "This application informed me of my music taste in a meaningful way."*

	Valid	Mean	Std. Deviation	Minimum	Maximum
Overall Musical Preferences	20	6.300	0.979	4.000	7.000
Informed Meaningfully	20	6.450	0.759	5.000	7.000

In addition to the written open-ended responses, participants also shared their thoughts while interacting with the experimenter. Among these comments, we would like to highlight the following: "Did you feed my survey responses to the assistant before we started this? It's as if it's referring to questions I answered in the same way that I answered them", "When I asked for some recommendations based off of the songs that I uploaded, the first song it recommended was one I almost chose for my favorites, and it isn't a very well-known song", "I'm older, so many of the songs I chose are from the 70s and 80s. When I first asked it for recommendations based on my song choices, it recommended some of my other favorite songs. However, what was truly amazing was that when I asked it for recommendations from modern day artists, it gave a selection of songs from some of my son's favorite artists. It was really amazing to see the influence my listening habits had on him, and how we can both appreciate the modern day renditions of the music of my generation."

## V. DISCUSSION

The *SoundSignature* application we present advances the state-of-the-art in the IoS by bridging MIR and semantic audio with AI and NLP, using cloud computing to provide users with personalized insights into their musical preferences in an innovative way. The results of the pilot user study suggest that the application was effective and highlights its educational

potential and novelty within the MIR landscape. The results also underscore the application's ability to deliver meaningful and personalized insights into users' musical preferences. Participants agreed with the conclusion provided by the app regarding their favorite music and also felt that the app informed them about their musical taste in a meaningful way (in both cases the average rating agreement was above 6 in a scale whereby 7 means "Completely Agree"). Results thus suggest that the application effectively bridged the gap between complex music analysis and user-friendly interaction. Furthermore, several participants were excited about questions they could ask following the original output (e.g., "What do you think that the songs I chose say about me as a person?", "What careers do you think I might excel in based on my musical taste and the characteristics of my favorite songs?"; See Supplementary Materials for examples).

We hypothesize that the lyrical analysis and cultural and historical context provided by the Open AI assistant provided users with a holistic understanding of their musical taste. Instead of simply being told that they appreciate bright, high-energy songs in major keys, the assistant can provide reports such as "Your choices reflect a blend of introspection, social awareness, and cultural pride, suggesting that you value both personal and collective narratives in music." This approach aligns with contemporary research emphasizing the importance of context and personalization in music recommendation systems [21, 43]. By contextualizing the lyrical themes and the overlapping characteristics of the songs, *SoundSignature* might be able to offer a more personal and engaging user experience.

*SoundSignature* aims to instill a sense of confidence and control in the users, allowing them to manipulate and understand their favorite songs by simply interacting with the chatbot, echoing the findings of prior research which emphasizes the importance of interactive learning tools in music education. The integration of these features makes *SoundSignature* not only a tool for entertainment and personal use but also a powerful educational resource. Indeed, we believe that the educational impact of *SoundSignature* is one of its most noteworthy potential contributions. By translating complex musical data into understandable insights, the application aligns with educational theories that advocate using technology to make learning more accessible and engaging [44] and with the increased demand to include semantic audio features in IoT devices [22]. Users reported that the application provided them with information about their musical preferences and helped them understand the underlying musical concepts. A tool that provides insights into not only *what* type of music a person likes but also *why* they might like it has the potential to help clinical populations engaged in music-based interventions (MBIs) [45]. Indeed, MBIs can benefit from personalized and enriched music delivery systems where the stimuli used during therapy is individualized to the patient's preferences, thus potentially maximizing the effectiveness of the therapy provided [46].

This educational aspect will be essential in future versions of the app, as it can empower users with knowledge traditionally restricted to academic or professional settings, thereby fostering a more informed and engaged base of music listeners. This democratization of music analysis tools aligns with current pedagogical approaches where music serves as a contextual tool

to facilitate deeper learning [16]. We hope the app will challenge the boundaries of traditional MIR systems by placing its primary focus on the user and by combining technical music analysis, educational content, and user-centric interfaces to cater to a diverse audience ranging from casual listeners to music enthusiasts and professionals [20].

We believe that *SoundSignature's* capability extends beyond personal use: it might also serve as a tool for building extensive music databases. The application can facilitate the creation of a rich music database by storing the musical and acoustic features alongside the metadata about the artists and lyrics it extracts. A scaled version of the app that allows users to create accounts and create a personal database could support more nuanced music recommendations that are based on acoustical and musical properties, but also on mood, themes, and even lyrical content, going beyond traditional recommendation algorithms. This capability could be especially beneficial if users started giving the chatbot more feedback in their responses or if the app offered "thumbs-up" and "thumbs-down" responses. Then, users could tell the chatbot which recommendations they like, whether something was inaccurate in the output, etc. Moreover, if we allow users to connect with their friends on the app, we could introduce shared reports that compare and contrast their *SoundSignature* profiles. This feature would foster more interactivity between users, enabling social engagement through music. For instance, users could explore shared song recommendations or create collaborative playlists based on their combined music profiles. This would align with the broader vision of the IoT and IoT-MT, where interconnected users and devices collaborate to enhance creative exploration and personalization [22, 47]. By integrating social elements into the app, *SoundSignature* could leverage the IoT ecosystem to provide a more dynamic and interactive user experience, influencing how users engage with music and with each other.

Similarly, this application could label data for music datasets used for deep learning. Generative applications like *Riffusion* [48] and *MusicalLM* [49] use diffusion models to create music from text inputs and are trained on spectrograms along with their labels. *SoundSignature* offers a unique opportunity to enrich these datasets by capitalizing on the feature set it extracts. By associating natural language descriptions and detailed musical and acoustic analyses with traditional spectrogram data, *SoundSignature* could provide richer, more contextual labels for training deep learning models for music generation. Being trained on a dataset with enriched labels could enhance a model's ability to generate music that resonates more deeply with human emotional and aesthetic preferences. Picture a model that could generate new compositions by interpolating and extrapolating from the features and styles of music the user prefers. For instance, users could upload their favorite songs, and *SoundSignature* could use the aggregated data to generate a new song. This song could mimic the input tracks' style, texture, and chords and attempt to incorporate the underlying emotional elements that the analysis has identified as significant to the user. Such a capability would be an exciting addition to the personalized music technology landscape, moving from passive analysis to active creation of user-specific content. This addition would also align with emerging trends in AI, where the focus is shifting towards more personalized, context-aware applications

[21, 43]. In addition, creating datasets that include enriched labels could also be useful for IoT devices and ecosystems, as they strongly depend on databases that include sound and music-related information to function [22]. The foundation laid by *SoundSignature's* data analysis capabilities provides a promising starting point for such developments.

#### A. Limitations

While the application has shown promising results, several limitations must be addressed. First, on closer inspection of the tempo analysis for the songs provided by participants in the pilot study, we realized that the tempo detection algorithm employed using *librosa* sometimes miscalculated BPM in half-time or double-time, which can misrepresent the song's actual tempo. This error can cause the chatbot to refer to a song as slow and laid-back when it might actually be upbeat and energetic, or vice versa. While there are other machine learning tools to estimate tempo (i.e., *madmom*) [50], these algorithms are computationally expensive and take longer to run, which can be a caveat and decrease the interactivity and ease-of-use of the app.

Additionally, the initial responses from the chatbot can sometimes be generic, as mentioned by one participant in the open-ended responses. However, when users engage with the system by asking more detailed questions (e.g., "What do these song choices say about me as a person?") the responses become significantly more personalized and insightful (see the qualitative results section). This indicates that while the chatbot is capable of deeper analyses, its effectiveness in providing the user with deeper details or ideas it has formed is contingent on user interaction, suggesting a need for more intuitive query handling and response generation.

Another limitation is that the OpenAI model the custom assistant runs on is only trained on data until October 2023. While this still allows for millions of songs and thousands of bands to be accounted for, it also means that the assistant may have difficulties providing details for newer artists. Interestingly, though, when conducting initial experiments, we uploaded an unreleased song by an unknown artist, and the chatbot was still able to provide interesting insights about the song and could somehow infer which genre it was (progressive rock) based on the features it was fed. Thus, the sections of the output that this knowledge cutoff date affects are the Cultural and Historical Context and Lyrical Content sections.

A notable limitation of our study is the predominance of participants with a Western musical enculturation and the selection of primarily Western songs. This bias may limit the generalizability of our findings, as musical preferences and perceptions can vary significantly across different cultures and musical traditions. While it did still seem to accurately report on the non-Western songs that were chosen, the insights provided by *SoundSignature* may not fully capture the diversity of musical tastes and preferences present in non-Western contexts. Future studies should aim to include a broader range of musical styles and participants from diverse cultural backgrounds to enhance the applicability and inclusivity of the application's analyses and recommendations. Finally, future user studies should include a control condition to ensure that the conclusions provided by the app are actually specific to the participant (e.g.,

provide participants with two app generated summaries of their musical taste, one stemming from the songs they selected and another from songs selected by another participant) [51, 52].

## VI. FUTURE DIRECTIONS

Looking forward, several enhancements are planned for *SoundSignature*. We chose a number of features to be extracted capitalizing on previous MIR, music psychology, and music neuroscience literature [30-32]. Nonetheless, there are many other available acoustic and music features that can be extracted (e.g., Attack Time, Roughness, Fluctuation Centroid and Entropy, etc.) and added to the pipeline. While we can infer from the user study that users are satisfied with the current state of the application, we plan to carry out further A/B testing to see whether people prefer a different combination of features or different iterations of the prompt that we provide the OpenAI Assistant with. Incorporating additional features may or may not enrich the data fed to the chatbot, and slight alterations to the prompt might significantly change how the assistant responds.

We also plan to conduct user studies with artists, composers, and producers who could provide insights into the app's utility for professional users with advanced musical knowledge, helping refine the tool's accuracy and user interface. This would also allow us to understand whether the application is meeting the current demands of the music industry. A music producer might utilize the source separator to sample the drums or vocals for a new beat. In contrast, a classical composer may first utilize the source separator to extract the strings section of a piece and then run the audio-to-midi on that stem to notate the piece and conduct a harmonic analysis. We want to ensure that advanced questions and function calls are accurate, run smoothly, and meet the needs of potential users.

Eventually, we would also like to add measures related to cognitive processing to *SoundSignature's* analysis pipeline, such as musical surprisal. Over the past several years, research on musical complexity and surprisal in cognitive neuroscience and music psychology has gained traction, leading to computational models that calculate prediction errors from music (e.g., *IDyOM*) [53]. The concept of musical surprise ties into theories proposing that an inverted U-shaped relationship exists between complexity and listener preference [54, 55]. These theories suggest that there is an optimal level of complexity that is variable for each individual depending on their musical experiences, musical training, etc. This variability implies that each listener has a unique internal model of what distinguishes something as complex, where one person's sweet spot might be boring or overwhelming to another. Thus, a future version of the app could include a measure of musical complexity. For example, we plan to include the Dynamic Regularity Extraction model (D-REX) to the app's pipeline [56]. D-REX is a Bayesian model of auditory salience that processes auditory features such as energy, pitch, and temporal modulation of continuous sounds to calculate surprisal scores, factoring in cognitive constraints like finite working memory and observation noise [56]. While *IDyOM* is similar in its approach to modeling musical surprisal, it only works with MIDI data, and thus cannot be applied to full songs [53]. Furthermore, recent findings from our lab show that the D-REX output correlates with both behavioral and neural measures of musical surprise



[57, 58]. In other words, the model ratings of surprisal can be used as a proxy for musical complexity. The addition of the D-REX measures would add quite an interesting dimension to the report (e.g., “Your music choices reflect a preference for music that is unpredictable and complex.”).

Moreover, integrating personality assessments could provide fascinating insights into the relationships between musical preferences, personality traits, and acoustic features, broadening the scope of personalized music recommendations and psychological studies in music perception [59-65]. Note that several participants showed an interest in getting information about how the output of *SoundSignature* relates to their personality traits and/or mood state (see qualitative results section). Much music preference research has examined the relationship between personality and music preference. However, many findings are incongruent and merely focus on associating broad personality traits with preferences for specific genres [59]. Interestingly, research has consistently demonstrated that specific musical features elicit distinct patterns of neural responses (fMRI) in participants with a similar musical enculturation. For example, timbral features have been associated with activations in cognitive areas of the cerebellum and sensory networks, while the processing of musical pulse and tonality recruits both cortical and subcortical circuits involved in cognitive, motor, and emotion-related functions [32, 66]. Such findings suggest there may be a neural basis for the interaction between acoustic features and emotional and cognitive processing during music listening, at least within participants with the same musical enculturation.

## VII. CONCLUSION

The results from the user study suggest that *SoundSignature* effectively fulfills its objectives of providing deep, personalized insights, and showcases the app’s potential to serve as an educational platform. These findings lay a promising foundation for future enhancements and the wider application of the technology. By addressing its current limitations and exploring new avenues for user engagement, *SoundSignature* is well-positioned to become a key tool for music lovers and professionals alike, reshaping the landscape of music analysis and appreciation for the general public including clinical populations [45].

As part of the broader IoS and IoMusT ecosystems, *SoundSignature* contributes to the development of interconnected musical tools and platforms that allow users to engage with music in new, interactive ways. By leveraging cloud-based processing and offering intuitive access to advanced music analysis tools without requiring technical expertise, the app aligns with the IoS vision of a networked, interactive audio landscape that includes semantic audio features. Through future expansions, such as integrating real-time interactivity and shared user reports, *SoundSignature* can further enhance social connectivity and creative exploration through music, offering new possibilities for personalized, context-aware music experiences and enrich current music and sound-related databases. As a result, it not only advances the state of music technology but also contributes to the evolving IoS framework, making sophisticated music analysis tools accessible to a wider audience.

## ACKNOWLEDGMENTS

Research reported in this publication was supported by the NCCIH of the National Institutes of Health under award number R34AT012943.

## REFERENCES

- [1] M. J. Walsh, “‘I’m also slightly conscious of how much I’m listening to something’: Music streaming and the transformation of music listening,” *Media, Culture & Society*, vol. 46, no. 4, pp. 745-761, 2024, doi: 10.1177/01634437231219142.
- [2] “Year on TikTok 2021 - Music Report.” TikTok. <https://newsroom.tiktok.com/en-us/year-on-tiktok-music-report-2021> (accessed 2022).
- [3] *Engaging with Music*. (2021). London, United Kingdom: International Federation of Phonographic Industry. [Online]. Available: <https://www.ifpi.org/wp-content/uploads/2021/10/IFPI-Engaging-with-Music-report.pdf>
- [4] “Thanks a billion!” TikTok. <https://newsroom.tiktok.com/en-us/1-billion-people-on-tiktok> (accessed).
- [5] K. R. M. Sanfilippo, N. Spiro, M. Molina-Solana, and A. Lamont, “Do the shuffle: Exploring reasons for music listening through shuffled play,” (in eng), *PLoS One*, vol. 15, no. 2, p. e0228457, 2020, doi: 10.1371/journal.pone.0228457.
- [6] M. L. Barata and P. S. Coelho, “Music streaming services: understanding the drivers of customer purchase and intention to recommend,” (in eng), *Heliyon*, vol. 7, no. 8, p. e07783, Aug 2021, doi: 10.1016/j.heliyon.2021.e07783.
- [7] I. Kamekhosh, G. Bonnin, and D. Jannach, “Effects of recommendations on the playlist creation behavior of users,” *User Modeling and User-Adapted Interaction*, vol. 30, no. 2, pp. 285-322, 2020/04/01 2020, doi: 10.1007/s11257-019-09237-4.
- [8] M. Furini and F. Fragnelli, “Social music discovery: an ethical recommendation system based on friend’s preferred songs,” *Multimedia Tools and Applications*, 2024/06/11 2024, doi: 10.1007/s11042-024-19505-0.
- [9] K. Dinnissen and C. Bauer, “Fairness in Music Recommender Systems: A Stakeholder-Centered Mini Review,” (in eng), *Front Big Data*, vol. 5, p. 913608, 2022, doi: 10.3389/fdata.2022.913608.
- [10] P. Bello and D. Garcia, “Cultural Divergence in popular music: the increasing diversity of music consumption on Spotify across countries,” *Humanities and Social Sciences Communications*, vol. 8, no. 1, p. 182, 2021/07/27 2021, doi: 10.1057/s41599-021-00855-1.
- [11] R. Prey, “Nothing personal: algorithmic individuation on music streaming platforms,” *Media, Culture & Society*, vol. 40, no. 7, pp. 1086-1100, 2018, doi: 10.1177/0163443717745147.
- [12] B. J. Hraes and J. Webster, “From selling songs to engineering experiences: exploring the competitive strategies of music streaming platforms,” *Journal of Cultural Economy*, vol. 14, no. 2, pp. 240-257, 2021/03/04 2021, doi: 10.1080/17530350.2020.1819374.
- [13] T. Schäfer, P. Sedlmeier, C. Städtler, and D. Huron, “The psychological functions of music listening,” (in eng), *Front Psychol*, vol. 4, p. 511, 2013, doi: 10.3389/fpsyg.2013.00511.
- [14] L. Ferreri, N. Singer, M. McPhee, P. Ripollés, R. J. Zatorre, and E. Mas-Herrero, “Engagement in Music-Related Activities During the COVID-19 Pandemic as a Mirror of Individual Differences in Musical Reward and Coping Strategies,” (in English), *Frontiers in Psychology*, Original Research vol. 12, 2021-June-28 2021, doi: 10.3389/fpsyg.2021.673772.
- [15] M. Schedl, E. Gómez, and J. Urbano, “Music Information Retrieval: Recent Developments and Applications,” *Foundations and Trends® in Information Retrieval*, vol. 8, no. 2-3, pp. 127-261, 2014, doi: 10.1561/15000000042.
- [16] M. Mueller, B. McFee, and K. Kinnaid, “Interactive Learning of Signal Processing Through Music: Making Fourier Analysis Concrete for Students,” *IEEE Signal Processing Magazine*, vol. 38, no. 3, pp. 73-84, 2021, doi: 10.1109/MSP.2021.3052181.



- [17] R. A. Wissner, "A Review of Public Musicology," *Journal of Musicological Research*, vol. 42, no. 4, pp. 228-231, 2023/10/02 2023, doi: 10.1080/01411896.2024.2336854.
- [18] B. Mor, S. Garhwal, and A. Kumar, "A Systematic Literature Review on Computational Musicology," *Archives of Computational Methods in Engineering*, vol. 27, no. 3, pp. 923-937, 2020/07/01 2020, doi: 10.1007/s11831-019-09337-9.
- [19] OpenAI. "OpenAI Assistants." OpenAI. <https://platform.openai.com/playground/assistants> (accessed 07/07, 2024).
- [20] M. Schedl, A. Flexer, and J. Urbano, "The neglected user in music information retrieval research," *Journal of Intelligent Information Systems*, vol. 41, no. 3, pp. 523-539, 2013/12/01 2013, doi: 10.1007/s10844-013-0247-6.
- [21] M. Fell, "Natural language processing for music information retrieval : deep analysis of lyrics structure and content  
Traitement automatique des langues pour la recherche d'information musicale : analyse profonde de la structure et du contenu des paroles de chansons," Université Côte d'Azur, 2020COAZ4017, 2020. [Online]. Available: <https://theses.hal.science/tel-02587910>
- [22] L. Turchet *et al.*, "The Internet of Sounds: Convergent Trends, Insights, and Future Directions," *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11264-11292, 2023, doi: 10.1109/JIOT.2023.3253602.
- [23] L. Turchet, C. Fischione, G. Essl, D. Keller, and M. Barthet, "Internet of Musical Things: Vision and Challenges," *IEEE Access*, vol. 6, pp. 61994-62017, 2018, doi: 10.1109/ACCESS.2018.2872625.
- [24] L. Turchet, G. Fazekas, M. Lagrange, H. S. Ghadikolaei, and C. Fischione, "The Internet of Audio Things: State of the Art, Vision, and Challenges," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10233-10249, 2020, doi: 10.1109/JIOT.2020.2997047.
- [25] L. Turchet, F. Antoniazzi, F. Viola, F. Giunchiglia, and G. Fazekas, "The Internet of Musical Things Ontology," *Journal of Web Semantics*, vol. 60, p. 100548, 2020/01/01/ 2020, doi: <https://doi.org/10.1016/j.websem.2020.100548>.
- [26] S.-W. Lee, B. Taylor, and G. Essl, "Interactive Music on the Web," in *Foundations in sound design for interactive media: A multidisciplinary approach*, M. Filimowicz Ed.: Routledge, 2019, pp. 200-228.
- [27] L. Turchet, J. Pauwels, C. Fischione, and G. Fazekas, "Cloud-smart Musical Instrument Interactions: Querying a Large Music Collection with a Smart Guitar," *ACM Trans. Internet Things*, vol. 1, no. 3, p. Article 15, 2020, doi: 10.1145/3377881.
- [28] B. McFee *et al.*, "librosa: Audio and Music Signal Analysis in Python," in *SciPy*, 2015, pp. 18-24, doi: 10.25080/Majora-7b98e3ed-003.
- [29] D. Bogdanov *et al.*, "ESSENTIA: an open-source library for sound and music analysis," presented at the Proceedings of the 21st ACM international conference on Multimedia, Barcelona, Spain, 2013. [Online]. Available: <https://doi.org/10.1145/2502081.2502229>.
- [30] I. Salakka *et al.*, "What makes music memorable? Relationships between acoustic musical features and music-evoked emotions and memories in older adults," (in English), *PLoS ONE*, vol. 16(5 May) (no pagination), May 2021, Art no. e0251692, doi: <https://dx.doi.org/10.1371/journal.pone.0251692>.
- [31] K. Groves, M. M. Farbood, P. Ripollés, and A. Zuanazzi, "Through the lens of music: Imagining movie scenes through soundtrack listening."
- [32] V. Alluri, P. Toivainen, I. P. Jääskeläinen, E. Glerean, M. Sams, and E. Brattico, "Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm," *NeuroImage*, vol. 59, no. 4, pp. 3677-3689, 2012/02/15/ 2012, doi: <https://doi.org/10.1016/j.neuroimage.2011.11.019>.
- [33] G. Fazekas and T. Wilmering, "Semantic web and semantic audio technologies," in *Tutorial presented at the 132nd Convention of the Audio Engineering Society, Budapest, Hungary*, 2012.
- [34] S. Rouard, F. Massa, and A. Defossez, "Hybrid Transformers for Music Source Separation," presented at the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023.
- [35] B. McFee and J. P. Bello, "Structured Training for Large-Vocabulary Chord Recognition," in *International Society for Music Information Retrieval*, 2017, pp. 188-194.
- [36] R. M. Bittner, J. J. Bosch, D. Rubinstein, G. Meseguer-Brocal, and S. Ewert, "A lightweight instrument-agnostic model for polyphonic note transcription and multipitch estimation," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022: IEEE, pp. 781-785.
- [37] M. Khorasani, M. Abdou, and J. Hernández Fernández, "Web Application Development with Streamlit," *Software Development*, pp. 498-507, 2022.
- [38] E. Mas-Herrero, J. Marco-Pallares, U. Lorenzo-Seva, R. J. Zatorre, and A. Rodríguez-Fornells, "Individual Differences in Music Reward Experiences," *Music Perception*, vol. 31, no. 2, pp. 118-138, 2013, doi: 10.1525/mp.2013.31.2.118.
- [39] N. Martínez-Molina, E. Mas-Herrero, A. Rodríguez-Fornells, R. J. Zatorre, and J. Marco-Pallares, "Neural correlates of specific musical anhedonia," *Proceedings of the National Academy of Sciences*, vol. 113, no. 46, pp. E7337-E7345, 2016, doi: 10.1073/pnas.1611211113.
- [40] E. Mas-Herrero, N. Singer, L. Ferreri, M. McPhee, R. J. Zatorre, and P. Ripollés, "Music engagement is negatively correlated with depressive symptoms during the COVID-19 pandemic via reward-related mechanisms," *Annals of the New York Academy of Sciences*, vol. 1519, no. 1, pp. 186-198, 2023, doi: <https://doi.org/10.1111/nyas.14931>.
- [41] D. Müllensiefen, B. Gingras, J. Musil, and L. Stewart, "The Musicality of Non-Musicians: An Index for Assessing Musical Sophistication in the General Population," *PLOS ONE*, vol. 9, no. 2, p. e89642, 2014, doi: 10.1371/journal.pone.0089642.
- [42] Qualtrics. "Qualtrics." <https://www.qualtrics.com> (accessed 2024).
- [43] Y. Deldjoo, M. Schedl, and P. Knees, "Content-driven music recommendation: Evolution, state of the art, and challenges," *Computer Science Review*, vol. 51, p. 100618, 2024/02/01/ 2024, doi: <https://doi.org/10.1016/j.cosrev.2024.100618>.
- [44] C. Dittmar, E. Cano, J. Abeßer, and S. Grollmisch, "Music information retrieval meets music education," in *Multimodal Music Processing*, (Dagstuhl Follow-Ups, Volume 3. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Germany: Dagstuhl Publishing, 2012, pp. 95-120.
- [45] A. J. Sihvonen, T. Särkämö, V. Leo, M. Tervaniemi, E. Altenmüller, and S. Soinila, "Music-based interventions in neurological rehabilitation," *The Lancet Neurology*, vol. 16, no. 8, pp. 648-660, 2017/08/01/ 2017, doi: [https://doi.org/10.1016/S1474-4422\(17\)30168-0](https://doi.org/10.1016/S1474-4422(17)30168-0).
- [46] E. Edwards *et al.*, "NIH Music-Based Intervention Toolkit: Music-Based Interventions for Brain Disorders of Aging," (in eng), *Neurology*, vol. 100, no. 18, pp. 868-878, May 2 2023, doi: 10.1212/wnl.00000000000206797.
- [47] L. Turchet and F. Antoniazzi, "Semantic Web of Musical Things: Achieving interoperability in the Internet of Musical Things," *Journal of Web Semantics*, vol. 75, p. 100758, 2023/01/01/ 2023, doi: <https://doi.org/10.1016/j.websem.2022.100758>.
- [48] S. Forsgren and H. Martiros. "Riffusion - Stable diffusion for real-time music generation." <https://riffusion.com/about> (accessed).
- [49] A. Agostinelli *et al.*, "MusicLM: Generating Music From Text," *ArXiv*, cs.SD 2023, doi: 2301.11325.
- [50] S. Böck, F. Korzeniowski, J. Schlüter, F. Krebs, and G. Widmer, "Madmom: A new python audio and music signal processing library," in *Proceedings of the 24th ACM international conference on Multimedia*, 2016, pp. 1174-1178.
- [51] F. Curzel, F. Osiurak, E. Trän, B. Tillmann, P. Ripollés, and L. Ferreri, "Enhancing musical pleasure through shared musical experience," *iScience*, vol. 27, no. 6, 2024, doi: 10.1016/j.isci.2024.109964.
- [52] A. J. Blood and R. J. Zatorre, "Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion," *Proceedings of the National Academy of Sciences*, vol. 98, no. 20, pp. 11818-11823, 2001, doi: 10.1073/pnas.191355898.
- [53] M. T. Pearce, "Statistical learning and probabilistic prediction in music cognition: mechanisms of stylistic enculturation," *Annals of*

- the New York Academy of Sciences, vol. 1423, no. 1, pp. 378-395, 2018, doi: <https://doi.org/10.1111/nyas.13654>.
- [54] B. P. Gold, M. T. Pearce, E. Mas-Herrero, A. Dagher, and R. J. Zatorre, "Predictability and Uncertainty in the Pleasure of Music: A Reward for Learning?," *The Journal of Neuroscience*, vol. 39, no. 47, pp. 9397-9409, 2019, doi: 10.1523/jneurosci.0428-19.2019.
- [55] V. K. M. Cheung, P. M. C. Harrison, L. Meyer, M. T. Pearce, J.-D. Haynes, and S. Koelsch, "Uncertainty and Surprise Jointly Predict Musical Pleasure and Amygdala, Hippocampus, and Auditory Cortex Activity," *Current Biology*, vol. 29, no. 23, pp. 4084-4092.e4, 2019/12/02/ 2019, doi: <https://doi.org/10.1016/j.cub.2019.09.067>.
- [56] B. Skeritt-Davis and M. Elhilali, "A Model for Statistical Regularity Extraction from Dynamic Sounds," (in eng), *Acta Acust United Acust*, vol. 105, no. 1, pp. 1-4, Jan-Feb 2019, doi: 10.3813/aaa.919279.
- [57] E. B. Abrams, E. Muñoz Vidal, C. Pelofi, and P. Ripollés, "Retrieving musical information from neural data: how cognitive features enrich acoustic ones," in *Ismir 2022 Hybrid Conference*, 2022.
- [58] E. B. Abrams, R. Namballa, R. He, D. Poeppel, and P. Ripollés, "Elevator music as a tool for the quantitative characterization of reward," *Annals of the New York Academy of Sciences*, 2024.
- [59] T. Schäfer and C. Mehlhorn, "Can personality traits predict musical style preferences? A meta-analysis," *Personality and Individual Differences*, vol. 116, pp. 265-273, 2017/10/01/ 2017, doi: <https://doi.org/10.1016/j.paid.2017.04.061>.
- [60] M. J. M. H. Delsing, T. F. M. ter Bogt, R. C. M. E. Engels, and W. H. J. Meeus, "Adolescents' music preferences and personality characteristics," *European Journal of Personality*, vol. 22, no. 2, pp. 109-130, 2008, doi: <https://doi.org/10.1002/per.665>.
- [61] K. A. Corrigan and E. G. Schellenberg, "Liking music: Genres, contextual factors, and individual differences," in *Art, Aesthetics, and the Brain*. Oxford: Oxford University Press, 2015.
- [62] T. Chamorro-Premuzic, P. Fagan, and A. Furnham, "Personality and uses of music as predictors of preferences for music consensually classified as happy, sad, complex, and social," *Psychology of Aesthetics, Creativity, and the Arts*, vol. 4, no. 4, pp. 205-213, 2010, doi: 10.1037/a0019210.
- [63] E. Carlson, P. Saari, B. Burger, and P. Toivainen, "Personality and musical preference using social-tagging in excerpt-selection," *Psychomusicology: Music, Mind, and Brain*, vol. 27, no. 3, pp. 203-212, 2017, doi: 10.1037/pmu0000183.
- [64] A. Bonneville-Roussy, P. J. Rentfrow, M. K. Xu, and J. Potter, "Music through the ages: Trends in musical engagement and preferences from adolescence through middle adulthood," (in eng), *J Pers Soc Psychol*, vol. 105, no. 4, pp. 703-17, Oct 2013, doi: 10.1037/a0033770.
- [65] D. Boer, R. Fischer, M. Strack, M. H. Bond, E. Lo, and J. Lam, "How shared preferences in music create bonds between people: values as the missing link," (in eng), *Pers Soc Psychol Bull*, vol. 37, no. 9, pp. 1159-71, Sep 2011, doi: 10.1177/0146167211407521.
- [66] V. Alluri *et al.*, "From Vivaldi to Beatles and back: Predicting lateralized brain responses to music," *NeuroImage*, vol. 83, pp. 627-636, 2013/12/01/ 2013, doi: <https://doi.org/10.1016/j.neuroimage.2013.06.064>.