Incremental and multimodal visualization of discographies: exploring the WASABI music knowledge base

Aline Menin, Michel Buffa, Maroua Tikat, Benjamin Molinet, Guillaume Pelerin, Laurent Pottier, Franck Michel, Marco Winckler

Incremental and multimodal visualization of discographies: exploring the WASABI music knowledge base

Aline Menin
Université Côte d'Azur, CNRS, Inria
Sophia Antipolis, France
aline.menin@inria.fr

Michel Buffa
Université Côte d'Azur, CNRS, Inria
Sophia Antipolis, France
michel.buffa@inria.fr

Maroua Tikat
Université Côte d'Azur, CNRS, Inria
Sophia Antipolis, France
maroua.tikat@inria.fr

Benjamin Molinet
Université Côte d'Azur, CNRS, Inria
Sophia Antipolis, France
benjamin.molinet@inria.fr

Guillaume Pellerin
IRCAM
France
guillaume.pellerin@ircam.fr

Laurent Pottier
Université Jean Monnet
Saint-Etienne, France
laurent.pottier@univ-st-etienne.fr

Franck Michel
Université Côte d'Azur, CNRS, Inria
Sophia Antipolis, France
franck.michel@inria.fr

Marco Winckler
Université Côte d'Azur, CNRS, Inria
Sophia Antipolis, France
marco.winckler@inria.fr

ABSTRACT

Collaborations between artists is very frequent in music industry and data analysis can help understand the impact of collaborative projects on artists' careers. In this paper, we propose an incremental and multimodal visualization to support the exploration of artists' discography and collaborations. Visualization techniques are used here to support the discovery of meaningful patterns and causal relationships. Our approach combines meta-data, graphics, interactive audio, and MIR analysis to explore the data using a Web browser. As a case of study, we explore the data from the WASABI knowledge graph, which describes the discography of more than 77,000 artists, representing over 200,000 albums and two million songs. We designed a web platform that requests metadata of any artist/band, parses the results, and displays the data through a visualization technique that combines stream-graphs, timelines and network representation. Audio thumbnailing techniques have been used to summarize an entire album or discography, which is used through a Web Audio augmented player integrated into the visualization to support auditory exploration of data. Advanced in-depth exploration of individual songs is supported through MIR analysis performed and displayed directly in the web browser. We demonstrate the use and feasibility of our method through a series of usage scenarios and validate its usefulness through expert judgment.

1. INTRODUCTION

Music data is increasingly available in digital form. Currently, there are several datasets available on the Web that describe musical content in various multimedia dimensions, such as lyrics, chords, sounds, metadata, etc. [8].Therefore, digital methods are especially important in musicology (the study of music as a branch of knowledge or field of research distinct from composition or performance [21]) for storing, structuring, and analyzing large amounts of music data available in digital form [29]. In this context, visualization techniques are suitable tools to facilitate the access to the data, while being capable of highlighting relationships between structural elements of music. A recent survey [18] illustrates the importance of visualization in musicology through an overview of the visualization techniques used to represent music-related data, including musical collections (albums, playlists, music archives), musical works, artists, or data issued from audio analysis (song structure, musical dimensions, etc.).

In this paper, we support the exploration of artists' discographies and collaborations. Artist collaborations in music tend to result in successful songs [10]. Thus, analysis of artists' collaboration data can help to understand the impact of these collaborative projects on artists' careers and their music style. Despite the importance and essential aspect of cultural context (e.g., release dates, collaborations, locations), the analysis of the timbre and annotations of the audio signal, as well as careful listening to the songs, is essential when studying the acoustic characteristics of songs [25]. Therefore, in this paper, we also explore the potential of the audio dimension to further support musicologists on their analysis, as well as to improve user perception through an auditory data exploration approach.

As a case of study, we explore the data from the WASABI

RDF knowledge graph (KG) [8], which gathers metadata describing over two million commercial songs (200K albums and 77K artists – mainly from pop/rock culture). It includes metadata about songs, albums, and artists (e.g., artists, discography, producers, dates, etc.) retrieved from multiple data sources on the Web. This data is linked to metadata extracted from NLP (Natural Language Processing) analysis of song lyrics and MIR (Music Information Retrieval) analysis of song audio content.

The WASABI dataset contains data on commercial music recordings from 1922 to 2022, but it does not provide explicit information about collaborations between artists and how they intersect over time. In this paper, we present a visualization tool to support the exploration of artists' discographies and collaborations: a social network of artists and bands derived from their recordings and mapped to the dimensions of *association* (type of contribution) and *time* (progression of the artist's career). The contributions in this article are as follows:

- a web-based interactive tool to visualize the discography of artists / groups and collaborations between musical artists and groups across time;

- an auditory exploration approach based on audio thumbnailing to support user perception throughout the visual exploration process;

- a direct link with external services that support further exploration of songs through MIR analysis.

The remainder of this document is organized as follows. Section 2 summarizes previous approaches to visualizing artist discography and collaborations, multimodal visualization, and audio thumbnailing techniques. Section 3 describes the WASABI dataset and the data model of our visualization tool. Section 4 describes our approach to a multimodal visualization that combines graphics, audio, and MIR analysis to support music artist data exploration. Section 5 describes a short evaluation conducted with an expert in musicology to assess the effectiveness and usefulness of our approach. Sections 6 and 7 conclude this work while discussing its limitations and potential future work.

2. RELATED WORK

2.1 Music datasets

The Million Song Dataset [6] is one of the outstanding music datasets describing one million contemporary popular songs through metadata extracted primarily from audio content analysis. It provides in addition metadata describing albums, artists, labels, tracks and relationships (among artists, albums, tracks, etc.). A recent addition in this domain is *the WASABI dataset*[1], which consists of more than 2 million commercial songs and holds a rich set of cultural metadata, about songs, albums and artists, which were retrieved from multiple sources on the web that it draws on by reusing their identifiers, or references via links. It also contains metadata extracted from the NLP analysis of song lyrics and from the MIR (Music Information Retrieval) analysis of song audio content, done during the project. The RDF representation of the data set is publicly available. We explore the data from this dataset in our case study.

[1]https://github.com/micbuffa/WasabiDataset

2.2 Visualization of artist discography and collaboration network

As shown in [18], various visualization techniques exist to explore music-related data, including song and instrument structure, performance analysis, emotions, similarity and listening statistics, and music alignment. Nonetheless, a few solutions have been proposed to explore the social network of musical artists, at least not in terms of collaborative composition and recordings. For instance, Miller and al. [23] propose a network graph to represent the relationships between musicians described in the Linked Jazz database, where typical relations are familial, academic, or work related. In addition to exploring collaborative productions, we are also interested in visualizing the progress of those collaborations together within the whole discography of artists.

MusicianMap [30] is an interactive web-based tool that enables the visualization of relationships between artists, including band membership, musical collaborations, and similarity of musical genre. This tool superposes two visualization techniques: a node-link diagram where nodes represent artists and links represent a relationship between them; and a timeline to show the evolution of those relationships over time, while providing time-based information (i.e., quantity of sales and recordings issued). A similar approach has been adopted to explore the recording sessions between jazz musicians over time [13], with the possibility of highlighting the strength of the collaborations at various points of time.

MusicLynx [1] is a music web application that supports the exploration of a graph of artist based on their similarities. Built from the linking of several free public data sources (including the "Million Song Dataset", DBPedia, MusicBrainz, and AcousticBrainz), it provides a multi-faceted navigation platform, offering a graph-based representation of the similarity links (modeled via audio descriptors of rhythm, tonality, and timbre) between artists. The web application also features a music player that streams short previews of the artist's top tracks (if available in the database of the streaming service).

As MusicianMap represents artist collaboration over time, we consider this tool the closets with respect to the approach presented in this paper. Nonetheless, our approach differs from the MusicianMap and other previous works by the fact that we support the exploration of artist discography based on different production types (composition, production, and performance) and through an auditory dimension, which allows the user to listen to what they see on the visualization. Furthermore, our multimodal and incremental visualization approach supports a better understanding of the collaborations between artists or between members of the same group on a recording as it represents collaboration links according to the different roles played by the artist on the collaborative recording (composer, producer, performer). The visualization approach leverage the *focus+context* interaction technique to provide data exploration in different granularity levels (at artist and song levels). The audio dimension has been incorporated, allowing, at the macro level, to perceive in an audible way the musical evolution of an artist over time or the influences of the different collaborators on the musical style of each one, while at the micro level it is possible to explore the audio signal of a song by performing MIR analysis (on demand) of the sound and by adding very precise annotations.

2.3 Timbre descriptors processing and access

The notion of timbre has been deeply modified by the evolution of recording studio techniques. For instance, there is an increasing use of DSP (Digital Signal Processing) effects by musicians. Ever more, electronic music uses sounds produced by synthesis or electroacoustic music, where composition is based on multi-category sounds. Thus, timbre can no longer be tied solely to the instrumental origin of sounds, but rather to integrated acoustic qualities [9]. Until recently, the only possible descriptions of sound were related to its spectrum and to sonograms and were mainly qualitative [26, 27, 7].

The acoustic characteristics of a sound are complex and it is difficult to relate them to the perception one may have of them. The quality and nature of the sound that we perceive depends on many interdependent factors that can be measured using a number of acoustic descriptors [19] each of which represents a certain quality of a sound. What we perceive is a combination of these different parameters, but studies to analyze and understand the link between the measured values of these parameters and the perception by a listener are still recent and very partial. The field is huge and still remains to be explored to better understand how the sounds of different musics resemble each other, to be able to establish classifications, or to finely analyze the music.

Among the descriptors that are used by many authors, the spectral centroid is the most widespread and is generally associated with the brilliance of the sound (energy distributed in the upper midrange or treble). Other statistical data extracted from the spectrum, such as the spectral slope, is also commonly used. The dynamic aspects of sound, at a time scale of a second, are still rather poorly described, even if there are some descriptors on this topic, such as the spectral flux [11]. Many studies are currently focused on the detection and recognition of musical instruments [20], on the separation of different audio tracks [15], on the isolation and erasure of the voice [16], and automatic detection of words [22]. These studies are usually based on machine learning methods. Few studies focus on the quantification and correlation of perceived sound parameters, especially in the case of electric or electronic music study. By studying the acoustic characteristics of music, we want to be able to make classifications and establish similarities between various pieces of music, based on the perception of timbre. The objective is to determine a kind of sound signature of the sound [24], on the scale of a few seconds, a duration that allows us, for example, to instantly recognize a group or a song that is familiar to us. By adding the visualization of the sound signature and audio descriptors associated with the presentation of songs of various artists, we aim to improve the perception of the sound quality of those songs.

3. DATA AND DATASET

Our visualization technique has been applied to data retrieved from the WASABI dataset, accessible through a REST API[2] and a public SPARQL endpoint[3]. Hereafter, we describe the processes of retrieving metadata about artist discography and collaborations and the process of extracting audio thumbnails from songs to support our multimodal exploration of data via visual and auditory approaches.

3.1 Artist discography and collaborations

The query presented in the Listing 1 is used to obtain the discography of a particular artist from the WASABI KG. The results contain the list of published songs (described by a title, an identifier, a release date), the album to which each song belongs, and the list of performers, authors (writers) and producers. Each artist is described by a name, a type (person or group), and an identifier. An artist can be associated with one or more songs, and a song can be associated with one or more artists. This association is defined by one or more contribution types (i.e. performer, producer, writer).

The data obtained from the data set is transformed to fit the data model used to create the visualizations, which corresponds to a multidimensional dynamic collaboration network. In the first layer of the network, the nodes represent the artists and the links between them are defined by the songs they have worked on together (produced, performed, or composed together). The second layer of the network describes these collaborations at a finer level: nodes represent songs that are linked to each other by multiple links (one per type of contribution of the involved artists). The data model describes an artist's collaboration network, as well as their entire discography (whether or not there was a collaboration) over time. The resulting output data is a JSON document containing a list of albums and songs per artist, a list of nodes, a list of links, and a dictionary of artists each containing attributes such as id, name, type, lifespan (birth/death for a person and begin/end for a group), and the list of members when artist is a group.

```
prefix dcterms: <http://purl.org/dc/terms/>
prefix foaf:    <http://xmlns.com/foaf/0.1/>
prefix schema:  <http://schema.org/>
prefix wsb:     <http://ns.inria.fr/wasabi/ontology/>

select distinct ?song ?songName ?songReleaseDate
?album ?albumReleaseDate
?performerURI ?type
?artist ?performers ?writers ?producers where {
  { select * where {
  ?song a wsb:Song ;
    dcterms:title ?songName ;
    schema:releaseDate ?date .
  ?song mo:performer ?performerURI .
  ?performerURI foaf:name ?artist;
    a ?type .
  filter(?artist == "[artist name]")
  optional {
    ?song schema:album [ dcterms:title ?album ;
      schema:releaseDate ?albumReleaseDate ]  } } }
  { select ?song
    (group_concat(distinct ?n ;
    separator = '--') as ?writers)
    where {
      optional { ?song schema:author ?n } } }
  { select ?song
    (group_concat(distinct ?n ;
    separator = '--') as ?producers)
    where {
      optional { ?song mo:producer ?n } } }
  { select ?song
    (group_concat(distinct ?n ;
    separator = '--') as ?performers)
    where {
      optional { ?song mo:performer ?a .
        ?a foaf:name ?n . } } }
}
```

Listing 1: SPARQL query used to retrieve an artist discography and collaboration data from the WASABI KG.
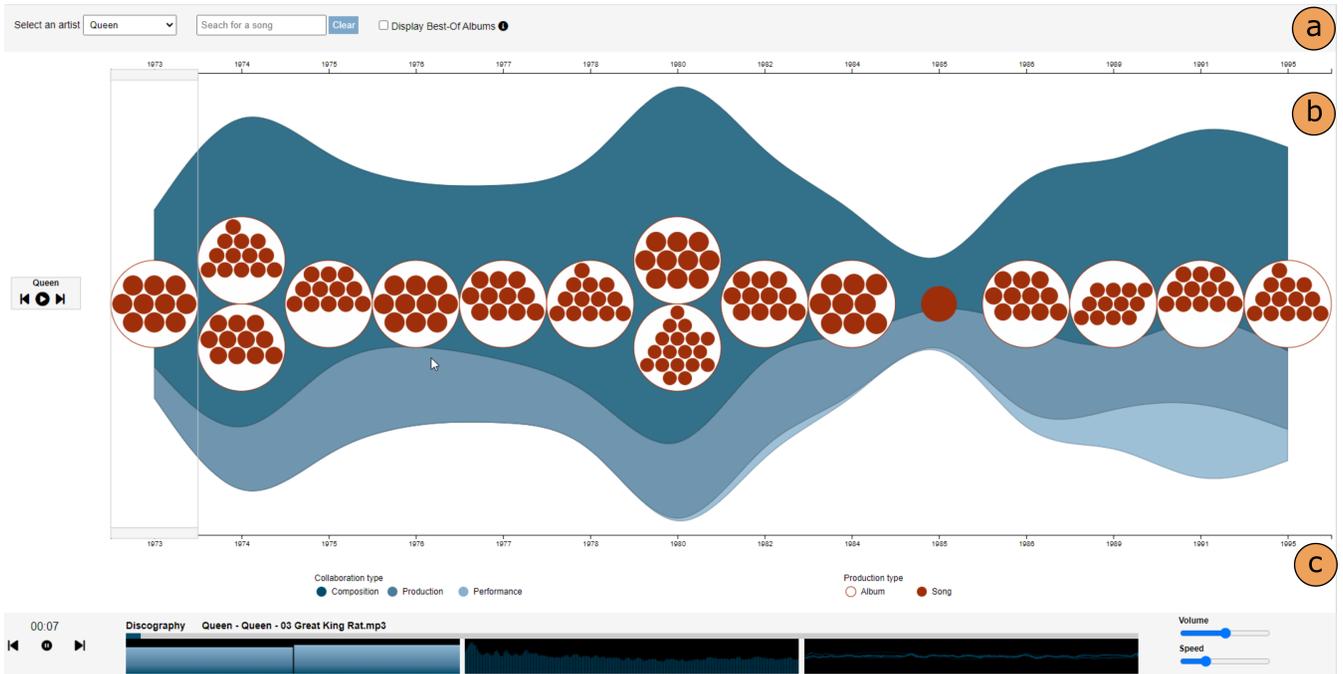
Figure 1: Overview of the visualization tool, including (a) a tool bar, where users can choose an artist to explore and search for particular songs in the visualization, (b) the visualization technique, displaying the discography of Queen band over time (from 1973 to 1995) as well as the songs of other artists in which they contributed, (c) an audio player, allowing the user to listen to a thumbnail of the artist discography or a particular album or yet, to listen to a whole song.

3.2 Audio thumbnails

To support auditory exploration of artist discography, we developed several scripts that take as input a hierarchy of audio files (discography/albums/songs) and produce as output audio summaries composed of selected excerpts of songs and metadata files. In our case, if a discography is composed of about ten music albums, each containing about ten songs, the final audio file will be made of about 100 song excerpts. If each excerpt lasts 3 seconds, i.e. 300 seconds or 5 minutes. The excerpts, ideally, should make the songs easily identifiable, especially for classic songs that have marked the history of music (by having been a number one hit for example). Several methods exist for extracting the most representative short part of a song in an unsupervised way, such as music summarization, audio thumbnailing, chorus detection, repeated pattern discovery, music structure analysis [17, 3, 4], and hybrid techniques combining song lyrics and audio content in one whole to get better results [12]. In its current proof-of-concept version, our solution simply select song extracts in the middle of the song, and shift a window backward in time to ensure that the full interval is not composed of silence.

A difficulty in associating audio albums with the WASABI dataset concerns first of all the notion of discography. In our case, after discussion with the musicologists we only considered studio albums in the original version of their release in their country of origin. For example, from the multiple versions of David Bowie's album "The Rise and Fall of Ziggy Stardust and the Spiders from Mars", the WASABI KG only describes the original version released in the UK in 1972, which included 11 songs. Thus, For a particular artist, we start with the audio albums in their original version, and

we organize the audio discography as a hierarchy of audio files with normalized file names. Through the exploration of these folders/files, for each album, we query the WASABI dataset via its REST API, and retrieve the WASABI id of the artist, albums, and songs. As we are comparing album/song names, the scripts include interactive hints when a proper match can not be achieved allowing to manually correct the information (e.g., it proposes suggestions when the exact name could not be found, or proposes to enter manually a new name). The metadata files also includes information such as the length of the excerpts used to build discography or album audio thumbnails. Both the *mp3* files containing these audio summaries and the metadata files are accessible via a single REST API.

4. INCREMENTAL AND MULTIMODAL VISUALIZATION APPROACH

In this visualization, we want to support the visual exploration of the following dimensions of the data: the *two-layer network* that describes the collaborations between artists (see the description above) and the *artist discography* that describes the albums and songs released over time. Furthermore, we seek to improve user perception by associating audio to the visualization, in a way that the users can listen to the songs visually represented in their screen. Hereafter, we describe how we combine graphics, audio, and MIR analyses into a multimodal visualization for artist discography and collaboration exploration.

The proposed visualization tool consists of three main elements: a tool bar (Fig. 1a), a visualization technique (Fig. 1b), and an audio player (Fig. 1c). In the tool bar,
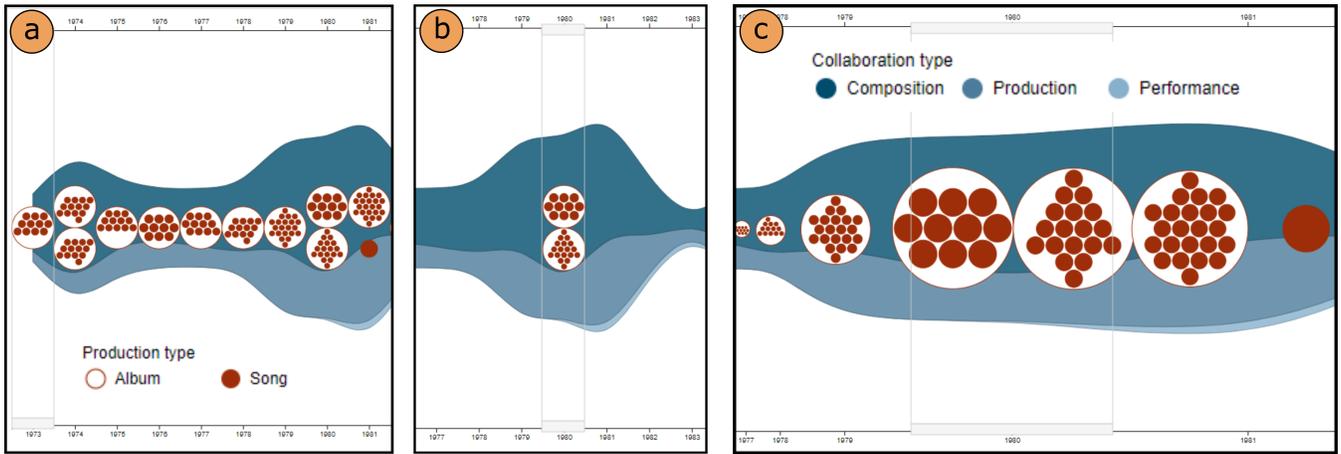
Figure 2: Focus+context exploration of artist discography. (a) Default presentation of albums and songs. (b) Focus on production of a particular year by dragging the slider over it. (c) Zoom into the productions of a particular year by clicking on the year of interest.

the user can select an artist to explore discography and collaboration data. Once the visualization is displayed, the user may also search for a particular song using the search bar. The visualization technique combines a timeline and a stream-graph to represent artist discography over time. The complete reasoning and features of the visualization technique are discussed in the following. Finally, the tool contains an audio player that allows users to listen to the songs displayed on the screen, as well as to a thumbnail of the artist discography or a particular album.

4.1 Visual exploration of discography

Typically, the time dimension can be mapped in an animation to a simulated time (time-to-time mapping) or to a space dimension of the generated visualization representing a timeline (time-to-space mapping) [5]. Timeline-based approaches provide a better overview of time, since they show the complete sequence of graphs in a static image [28], while animation-based approaches show one time slice at a time and, therefore, suffer with perception problems, as the user must rely on their own memory to understand the data and grasp the timely changes [2]. To support users to grasp at a glance the complete artist discography and to reduce cognitive effort, we chose a timeline-based visualization approach.

The proposed visualization follows the structure presented in Figure 1b. The name of the artist is placed on the left axis, whereas the top and bottom axes represent time in years, ranging from the release year of the first album to the last album. Albums are represented through a circle packing technique, where dark red circles represent the songs belonging to that album. Singles or songs from other artists in which the referred artist collaborated are represented as dark red circles placed on the side of albums. Notice that the label of the artist contains a player control, which serve to explore the artist discography by listening to a thumbnail of it (see more details below).

Furthermore, we represent the music production profile of an artist over time using a stream-graph technique, where each stream area represents the amount of contribution the artist had in the associated songs. We consider three types of contribution in a song (i.e., performance, production, or composition), which we color-coded through a blue gradient ranging from dark to light blue, respectively. The amount of production per year corresponds to the sum of songs released by the artist or to which the artist contributed. Through this visualization, the user can grasp rapidly when an artist was the most productive and what type of musical production they did.

We provide different levels of interaction to explore the discography of an artist (see Figure 2). The tool includes a slider that can be dragged over the x-axis to focus on the productions of different years, while progressively hiding the items that do not correspond to the selection. In a second level, the user can focus and zoom in on the productions of a particular year by selecting that year on the x axis (top or bottom), resulting in a distortion of these axes and expansion of items within the affected area (Figure 2c).

4.2 Auditory exploration of discography

At a third level of interaction, we allow the user to listen to the songs being represented. For that, an audio player using the WebAudio API has been specially developed to interact bidirectionally with the audio thumbnails. This player offers playback modes to explore the entire discography, an album, or a particular song. Written as a Web Component, it is integrated to the visualization tool and offers a complete JavaScript API. Thus, from the graphical visualization, it is possible to trigger the playback, to jump from one extract to another in audio summaries, to adjust the playback speed, etc. Particularly, the user can use the audio controls placed under the name of the artist or directly click on a circle representing either an album or a song to listen to the associated audio. Figure 3 shows the process to play the audio thumbnail of a particular album. When hovering over a particular album, the system displays a "play" symbol indicating that the user can listen to that album (Figure 3a). By clicking on it, the system launches the audio player for that particular album and goes into a state of "pause", as illustrated by the icon on the album (Figure 3b). The visualization shows the user which song thumbnail is currently playing by highlighting the circle that represents it. The audio player also shows this information by displaying the title

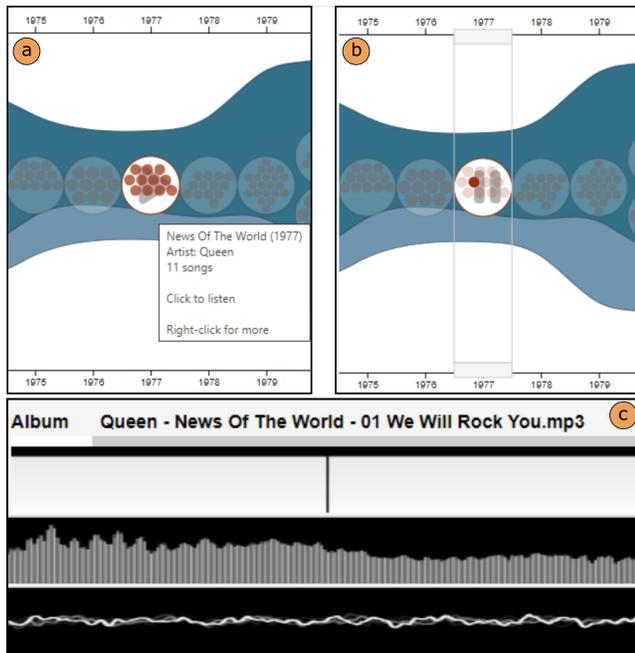of the song, while also showing the variations in balance and volume through the audio monitors (Figure 3c).

Figure 3: The interactive process that support listening to an album thumbnail. (a) By hovering an item, the "play" icon indicates that the user can listen to that item, while the tooltip shows information about the item. (b) By clicking on the item, the system launches the audio player for that item; the current played song is then highlighted in the visualization. (c) the audio player displays the name of the current song thumbnail and displays the changes in balance and volume accordingly.

4.3 Visual exploration of artist collaborations

We adopted an incremental approach to support the exploration of artist collaboration. Let us assume that we are interested in exploring the collaborations of Queen over time. When right-clicking on the artist's name, the system displays a context menu that provides a set of options for the user, including the list of artists with whom they collaborated throughout their career (Figure 4a). Upon selecting an artist from the list, let us say *Freddie Mercury*, the system queries the RDF graph to retrieve the data (albums, songs, and collaborators) corresponding to the selected artist and displays it as shown in Figure 4b. At this point, we are exploring data from two related artists, which means that our data correspond to the *two-layer network* mentioned above. Freddie Mercury, in this particular case, had a solo career, which solo albums can be seen on his timeline. However, we can also identify songs in which he participated, even if they are not in his solo albums. This is the case for all Queen songs he composed as a member of the group, represented by the dark red circles placed next to the albums (white circles).

Network data is typically visualized through node-link diagrams, where nodes are connected by links, or adjacency matrices, where vertices are mapped to rows and columns of the matrix, and a colored intersection cell encoding an edge [5]. As we are representing the time dimension using a timeline, we chose a node-link diagram to represent the collaborations between artists. In particular, we chose a juxtaposed layout to present the network and discography side by side. As mentioned above, our network consists of two layers: collaborations at the artist and song levels. Collaborations at the artist level are represented through single black lines that connect artists in each year where a collaboration has been detected. At this level, the user can grasp the information that a collaboration exists between those artists, without information regarding the type of collaboration. By hovering over an artist's label, the system highlights all the collaborative songs and albums of this artist, as well as the artists to whom those collaborations happened.

Exploring the second layer of the network is supported through a focus+context interaction technique. When clicking on a year of interest, the timeline is stretched horizontally, revealing the collaboration links between artists at the song level (Figure 4c). The visualization displays a link between each two nodes (per artist pair) representing the songs where both artists worked together. The links are colored to represent the different types of collaboration (performance, production, or composition) following the same color-code used to color the artist's temporal profile. The user can easily change the focus by clicking on another year on the top/bottom time axes or by dragging the slider over the axes to browse across the different years.

4.4 Visual exploration of timbre descriptors

In order to access to the timbre descriptors of each track described in 2.3, we support the analysis of audio signal by linking the WASABI dataset songs to a remote TimeSide audio processing service.

TimeSide is an open, scalable, audio processing framework implemented in Python that enables low- and high-level audio analysis, visualization, transcoding, streaming, and labeling. Initially designed for MIR and computational musicology usecases[14], it supports reproducible and extensible processing on large datasets of any audio or video format through state-of-the-art audio libraries (i.e. VAMP, Aubio, Essentia, Numpy/Scipy). To support the exploration of the WASABI songs, we have created or extended some parts of the framework such as a secured and documented REST API with JSON Web Token access capabilities, a Provider module to handle automatic extraction of YouTube and Deezer tracks or 30-second extracts, a JavaScript SDK[4] for the development of client applications, and a new web front-end prototype[5]. The Timeside remote service based at IRCAM is then used to run several timbre analyzers based on VAMP plugins[6] on demand. The WASABI server sends some metadata, the YouTube ID and the list of the needed processors in order to compute and store the result data of each track. As a result, the user of the WASABI main interface can explore the audio related data for a particular song from a context-menu embedded in each dark red circle that represents a song. The timbre descriptors curves are then displayed in an augmented multi-track audio player which provides a dynamic and zoomable interface to facilitate time based fine analyzing and labeling (Figure 5).
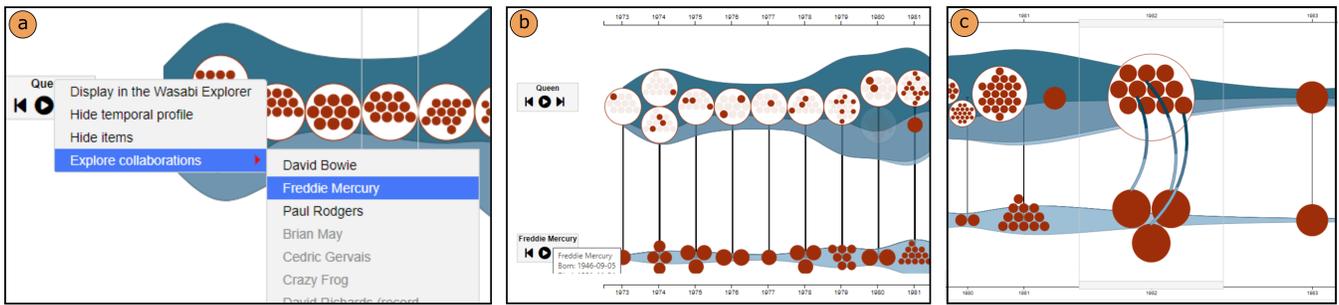

Figure 4: Exploration of artist collaboration. (a) Incremental integration of artist data by selecting a collaborator of the current artist. (b) Collaborations at the artist level represented by line segments between them. (c) Collaborations at the song level represented through colored arcs between them, where color encodes the type of contribution of each involved artist.
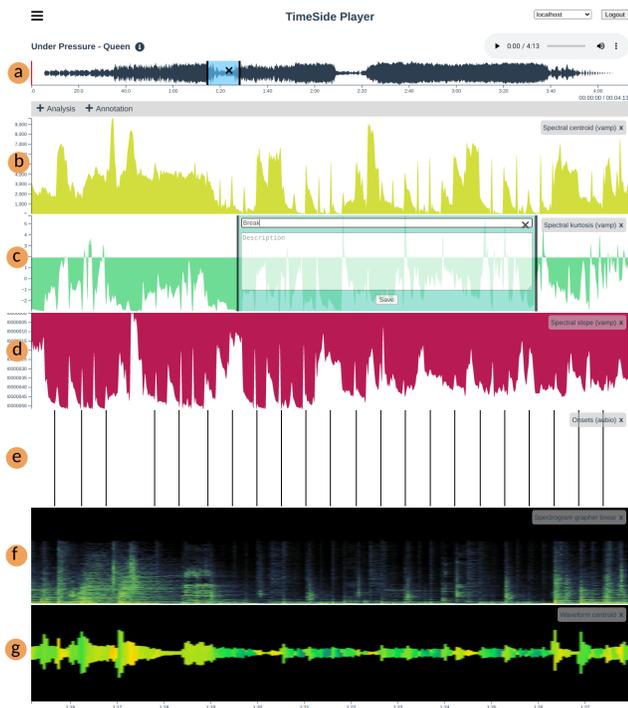

Figure 5: Overview of the Timeside-Player interface picturing timbre and rhythm analyses for the Queen song "Under Pressure". It shows (a) the audio player and the descriptors of (b) Spectral centroid, (c) Spectral kurtosis, (d) Spectral slope, (e) Onsets, (f) Linear Spectrogram and (g) Waveform with spectral centroid color from the time window selected on the first waveform. The selected segment on the (c) Spectral Kurtosis curve is used to annotate a sound event in this part.

In the same way, for the purpose of data check and to provide more information about the items to the user, our interface also directs the user to the source of data by right-clicking on the element of interest (i.e., the artist name, an album, or a song). As we can see in Figure 4a, the user can choose the option "Display in the Wasabi Explorer", which will redirect them to the artist's page in the WASABI Explorer[7].

_____

[7]https://wasabi.i3s.unice.fr/

5. FORMATIVE EVALUATION

We conducted a semi-structured interview of about one hour via a video conference call with an expert on musicology (over 15 years of experience), specialized in studying music through computer technology (e.g., computer-assisted composition, sound spatialization, timbre analysis, etc.). The goals of this interview were two-fold: (i) to better understand the needs of domain experts in terms of exploring artist discography and collaboration data and (ii) to assess the extent to which our visualization approach can support those needs.

5.1 Protocol and Questions

The interview followed a three-part protocol. It began with a set of questions regarding the user needs in terms of exploring artist discography and collaboration data. In particular, the interviewer asked them the following questions:

- Is exploring artist discography important? Why?

- When exploring discography, how important is to explore artist collaboration? Please provide examples.

- What is difficult when analyzing artist collaboration data?

- What attributes do you consider important when defining the discography/collaboration of two or more music artists? Rank them by importance

- In your opinion, how important is it to access audio information throughout the exploration of artist discography/collaboration?

In the second part of the interview, the interviewer showed the tool through a set of use case scenarios designed to guide the expert user during the semi-directed interview. These scenarios demonstrate the different aspects of the tool (described in Section 4), namely: (i) exploring the discography of a particular artist (Queen), (ii) using audio to explore the artist discography, (iii) exploring the collaborations of a particular artist, (iv) investigating the collaboration between two artists in a certain period of time, and (v) further exploring songs through external services (Wasabi Explorer and TimeSide). After each scenario, the interviewer asked them to identify (i) three things they like about the scenario, (ii) three things they dislike, and (iii) suggestions on how to improve it.

Finally, in the debriefing part, a set of general questions was asked about the usefulness of the tool to support their exploration needs. In particular, the interviewer asked them the following questions:

- In your opinion, what is the usefulness of a visualization technique [in general and this specific visualization] to explore artist discography?

- In your opinion, what is the usefulness of a visualization technique to explore artist collaboration?

- In your opinion, what is the usefulness of multimedia to explore artist discography and collaboration?

- Would you be willing to use this visualization? If yes, in what situation?

Throughout the interview, the questions could include follow-up questions based on the responses of the interviewee, which for the sake of simplicity is called domain expert.

5.2 Results

5.2.1 Exploration needs

According to our domain expert, exploring artist discography is important as it allows one to observe how the music of an artist evolves over time, as well as to identify the time periods when the artist were the most productive. In terms of artist collaboration, the domain expert had already studied the collaborations of an artist with sound engineers and compositors. When studying the collaborative projects between groups, it was considered important to identify the influences of a group on another in terms of style, sound and music diversity. However, collaboration information is not relevant for every artist, as not every collaborative project is interesting, i.e. not every collaborative project impacts the music style of an artist or results in successful songs.

When asked about the relevant attributes in artist discography, the domain expert mentioned music style of the artist/group, the variation of music styles over multiple pieces of the song (there is no taxonomy to describe this variation), tempo, and instrumentation. Regarding the relevant attributes in artist collaboration, the domain expert find important to provide elements that describe the influence of an artist over another; in particular, whether (or not) there were an influence over another artist and if that influence remains over the next albums. Another relevant aspect is to analyze the impacts the collaborative project had on the career of an artist and how that collaboration was received by the public. Finally, the domain expert believes that the usage of audio signal is fundamental when exploring artist discography and collaboration, as it allows one to illustrate the data (e.g., identify the influence of style by comparing audios).

5.2.2 Appraisal of the visualization tool

Overall, the domain expert was satisfied with the tool in all the different use case scenarios. The way we represented the discography of an artist and the easy and direct access to different information and tools was greatly appreciated.

In terms of the auditory exploration approach, the domain expert enjoyed the ability to click on a visual variable that represents a song and immediately listen to the audio extract while seeing the spectrum, volume, and balance variations

directly on the augmented audio player (a suggestion was made to make it larger to allow better perception of the visual variations). The domain expert also appreciated the ability to control the volume and speed of the audio and to play the audio summary of the previous or next album, which supports a dynamic browsing of the discography of artists.

When exploring collaborations between two artists over a certain period of time, the domain expert enjoyed the ability to zoom in on the productions and collaborations over a certain year while displaying the links between particular songs. It was mentioned that the number of links displayed could clutter the visualization, hindering the identification of the different colors encoding the different collaboration types. Nonetheless, the domain expert acknowledged that hovering an item or link highlights the important information, thus mitigating the cluttering effect.

In general, tool-tips displaying more information about an item was praised, but the domain expert would rather like to have more information (such as the history of the artist/group, the name of albums, etc.) displayed directly on the interface, reducing the mouse movements necessary to display more information about an item (e.g., a song). It was also suggested to have another visual representation when clicking on an album, which would enable the exploration of different types of songs.

The domain expert appreciated the link between the visualization tool and the external services; in particular, the link to the WASABI Explorer as it provides detailed information about the album, the song, or the artist to deepen the analysis. The domain expert was happy with the possibility of directly exploring the results of the MIR analysis on the audio signal of a particular song. However, the domain expert would prefer to have the TimeSide interface embedded in the visualization tool, instead of being redirected to a different window. The domain expert also proposed to include multiple ways of accessing those external services, such as through a button that opens up the WASABI Explorer or TimeSide without a particular song associated (which could be chosen directly in the respective interfaces).

A more general concern of the domain expert was related to the accuracy of the data, as the discography of an artist is sometimes represented with more albums then the actual discography (the original albums). The data on the Web, although powerful, is known to have many quality issues, which is one of the matters addressed by the participant. In the presented scenario (i.e. exploration of Queen discography), they observed that sometimes Queen does not appear as the composer of their own songs, particularly at the early years of their career. This particular case could be explained by the fact that, at the beginning, the songs were signed by whichever member of the group had composed them, it was only after 1995 that Queen started signing some of their songs as a band (i.e. Queen (band)). This is, however, a type of issue that requires knowledge about the data context to be fixed (or understood).

Finally, the domain expert said that they would use this visualization in their daily work for tasks such as performing research on artists, identifying important time periods of their productions, and exploring the discography of different artists.

6. DISCUSSION

The visualization approach proposed in this paper to assist in the exploration of artist discography and collaborations, raises some points for discussion, as follows:

Analytical tasks. We propose a four-fold approach to support the resolution of domain tasks. First, we propose a visual stream-graph representation that shows the variation of artist production over time together with their discography. Second, we combine multiple stream-graphs side-by-side to support (i) comparison of artist production over time and (ii) exploration of their interactions at different time periods, defined by different products (albums/songs). Third, we propose an auditory exploration of discographies, which allow users to listen to what they see, recognizing the data (successful songs) while increasing user perception. Finally, the fourth approach refers to the integration of external services such as TimeSide which allows to deepen the exploration of songs through MIR analysis.

Visual design and interactions. Our approach represents multiple and complex dimensions of the data: time, network, attributes. As we deal with musical data, we combine visual and auditory senses to support multiple domain-based tasks while improving user perception. We leverage a juxtaposed layout to provide clear representation of artist discography and their collaboration network side-by-side, while providing focus+context interaction to support a deeper exploration of collaboration data. To reduce visual clutter, our exploration approach also allows users to incrementally include or remove artist data, as well as to hide/show the products (albums/songs) or the production profile (stream-graph) according to the ongoing analysis and the space available. By clicking on artist, album, or song, the user can listen to an extract of the artist discography, album, or song, respectively. Furthermore, each artist, album, and song are linked to their respective page in the WASABI Explorer, and each song can be further explored on TimeSide through a simple selection on the visualization.

Auditory data exploration. To support auditory exploration of data, we extracted a set of audio thumbnails that represent the artist discography or a particular album and integrated them to the visualization. Currently, these audio thumbnails are generated through extracts of about 3 seconds from the middle of each song, using a shift function to avoid extracting silent audio. In the future work, as we have access to 1.7M of song lyrics in the WASABI dataset[8], we intend to use an hybrid technique based on song lyrics and audio content [12] to extract meaningful pieces of audio.

Usability and Suitability. We assessed our approach through a semi-structured interview with a musicologist. Although we have only interviewed one expert user, the results suggest that we have a promising approach to support musicologists in the resolution of domain-related tasks. Future work includes the design and implementation of user-based evaluations to investigate the usability of our approach for exploring artist discography and collaboration networks of different artists and its suitability for assisting the resolution of domain-related tasks.

Generalization. Although the use case study of this paper is the WASABI dataset, the visualization technique has been implemented in a way that it can be supplied with data describing the discography and collaboration of any artist regardless of the source. Moreover, the technique can be ex-

tended to explore other types of collaboration networks (e.g., scientific co-authorship). However, the auditory exploration feature relies on an audio thumbnailing procedure that is not currently automatized, as it requires a manual retrieval of songs that will feed the audio thumbnailing scripts. Future work includes to fully automatize this procedure to support a complete visual and auditory exploration of all music data.

7. CONCLUSIONS AND FUTURE WORK

We presented a web-based interactive tool that supports an incremental and multimodal exploration of music data through visual and auditory stimulus. In particular, our approach supports the analysis of artist discography and collaboration network. We have also linked our visualization with the TimeSide tool to support MIR analysis of the songs. We have demonstrated the feasibility and usefulness of our tool through a semi-structured interview with a domain (musicology) expert user, who gave us valuable feedback on the utility of our tool to solve the domain-related tasks and on the aspects that need improving. Future work includes the design and implementation of user-based evaluations to assess the usability and effectiveness of our approach to support domain-related tasks, as well as to support the exploration of data from different artists and sources. We provide auditory exploration of data through audio thumbnails played on-demand by the user while exploring the data in the visualization. Our proof-of-concept support auditory exploration of only a limited set of songs, which should be extended in future work. We also intend to improve the audio thumbnailing process by using the method proposed by Fell et al. [12] to enhance the exploration process by providing more meaningful audio extracts. The tool is available at http://dataviz.i3s.unice.fr/muvin.

8. ACKNOWLEDGMENTS

References

[1] A. Allik, F. Thalmann, and M. Sandler. Musiclynx: Exploring music through artist similarity graphs. In *Companion Proceedings of the The Web Conference 2018*, WWW '18, page 167–170, Republic and Canton of Geneva, CHE, 2018. International World Wide Web Conferences Steering Committee. ISBN 9781450356404. 10.1145/3184558.3186970.

[2] B. Bach, P. Dragicevic, D. Archambault, C. Hurter, and S. Carpendale. A Descriptive Framework for Temporal Data Visualizations Based on Generalized Space-Time Cubes. *Computer Graphics Forum*, 2016. 10.1111/cgf.12804.

[3] J. Balen. An overview of audio thumbnailing techniques and their relevance to the cogitch project. *Tech. Rep.*, 2012.

[4] M. Bartsch and G. Wakefield. Audio thumbnailing of popular music using chroma-based representations.

---

[8]Privately available (contact for potential academic collaboration)

IEEE Transactions on Multimedia, 7(1):96–104, 2005. 10.1109/TMM.2004.840597.

[5] F. Beck, M. Burch, S. Diehl, and D. Weiskopf. The State of the Art in Visualizing Dynamic Graphs. In R. Borgo, R. Maciejewski, and I. Viola, editors, *EuroVis - STARs*. The Eurographics Association, 2014. ISBN 978-3-03868-028-4. 10.2312/eurovisstar.20141174.

[6] T. Bertin-Mahieux, D. P. Ellis, B. Whitman, and P. Lamere. The million song dataset. In *Proc. of the ISMIR Conf.*, 2011. 10.7916/D8NZ8J07.

[7] P. Bootz and X. Hautbois. Les unités sémiotiques temporelles. In *Séminaire MaMuX, IRCAM*, 2009.

[8] M. Buffa, E. Cabrio, M. Fell, F. Gandon, A. Giboin, R. Hennequin, F. Michel, J. Pauwels, G. Pellerin, M. Tikat, et al. The wasabi dataset: Cultural, lyrics and audio analysis metadata about 2 million popular commercially released songs. In *European Semantic Web Conference*, pages 515–531. Springer, 2021.

[9] M. Chion. Dissolution of the Notion of Timbre. *differences*, 22(2-3):235–239, 12 2011. ISSN 1040-7391. 10.1215/10407391-1428906.

[10] A. Deshmane and V. Martinez-de Albeniz. Come Together, Right Now: An Empirical Study of Collaborations in the Music Industry. *SSRN Electronic Journal*, 2020. ISSN 1556-5068. 10.2139/ssrn.3743462.

[11] S. Dixon. Onset detection revisited. In *Proceedings of the 9th International Conference on Digital Audio Effects*, volume 120, pages 133–137. Citeseer, 2006.

[12] M. Fell, Y. Nechaev, G. Meseguer-Brocal, E. Cabrio, F. Gandon, and G. Peeters. Lyrics segmentation via bimodal text–audio representation. *Natural Language Engineering*, pages 1–20, 2021. 10.1017/S1351324921000024.

[13] D. Filippova, M. Fitzgerald, C. Kingsford, and F. Benadon. Dynamic exploration of recording sessions between jazz musicians over time. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing*, pages 368–376, 2012. 10.1109/SocialCom-PASSAT.2012.78.

[14] T. Fillon, J. Simonnot, M.-F. Mifune, S. Khoury, G. Pellerin, E. Amy de La Bretèque, D. Doukhan, D. Fourer, J.-L. Rouas, J. Pinquier, and C. Barras. Telemeta: An open-source web framework for ethnomusicological audio archives management and automatic analysis. In *1st International Digital Libraries for Musicology workshop*, 2014.

[15] R. Hennequin, A. Khlif, F. Voituret, and M. Moussallam. Spleeter: a fast and efficient music source separation tool with pre-trained models. *Journal of Open Source Software*, 5(50):2154, 2020. 10.21105/joss.02154.

[16] A. Jansson, E. Humphrey, N. Montecchio, R. Bittner, A. Kumar, and T. Weyde. Singing voice separation with deep u-net convolutional networks. Accessed on March 2022, October 2017. URL https://openaccess.city.ac.uk/id/eprint/19289/.

[17] N. Jiang and M. Müller. Towards efficient audio thumbnailing. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5192–5196, 2014. 10.1109/ICASSP.2014.6854593.

[18] R. Khulusi, J. Kusnick, C. Meinecke, C. Gillmann, J. Focht, and S. Jänicke. A survey on visualizations for musical data. *Computer Graphics Forum*, 39(6):82–110, 2020. 10.1111/cgf.13905.

[19] O. Lartillot, P. Toiviainen, and T. Eerola. Mirtoolbox 1.6. 1 user's manual. *Aalborg University, Denmark*, 2014.

[20] A. Livshin. *Automatic Musical Instrument Recognition and Related Topics*. PhD thesis, Université Pierre et Marie Curie-Paris VI, 2007.

[21] Merrian-Webster. Definition of musicology. https://www.merriam-webster.com/dictionary/musicology, 2018. Accessed in March 2022.

[22] A. Mesaros and T. Virtanen. Automatic recognition of lyrics in singing. *EURASIP Journal on Audio, Speech, and Music Processing*, 2010:1–11, 2010. 10.1155/2010/546047.

[23] M. Miller, J. Walloch, and M. C. Pattuelli. Visualizing linked Jazz: A web-based tool for social network analysis and exploration. *Proceedings of the American Society for Information Science and Technology*, 49(1):1–3, 2012. ISSN 1550-8390. 10.1002/meet.14504901295.

[24] L. Pottier. Local sound signatures for music recommendations. In *Proceedings of Timbre 2018: Timbre Is a Many-Splendored Thing, , Montréal, Québec, Canada*, pages 65–66, 2018.

[25] L. Pottier. Étude des caractéristiques acoustiques de quatre pièces du groupe led zeppelin et comparaison avec d'autres répertoires. In *Led Zeppelin – Contexte, analyse, réception*, pages 109–132. Ph. Gonin, Dijon, EUD, 2021.

[26] P. Schaeffer. *Traité des objets musicaux*. Média Diffusion, 2016.

[27] D. Smalley. Spectro-morphology and structuring processes. In *The language of electroacoustic music*, pages 61–93. Springer, 1986. 10.1007/978-1-349-18492-7_5.

[28] B. Tversky, J. B. Morrison, and M. Betrancourt. Animation: can it facilitate? *International journal of human-computer studies*, 57(4):247–262, 2002. 10.1006/ijhc.2002.1017.

[29] M. Urberg. Pasts and futures of digital humanities in musicology: Moving towards a "bigger tent". *Music Reference Services Quarterly*, 20(3-4):134–150, 2017. 10.1080/10588167.2017.1404301.

[30] J.-D. Yim, C. Shaw, and L. Bartram. Musician map: Visualizing music collaborations over time. In *Visualization and Data Analysis, VDA 2009*, page 72430, 2009. 10.1117/12.812377.