Open Your Ears and Take a Look:
A State-of-the-Art Report on the Integration of Sonification and Visualization

K. Enge[1,2] , E. Elmquist[3] , V. Caiola[4] , N. Rönnberg[3] , A. Rind[1] , M. Iber[1] ,
S. Lenzi[5,6] , F. Lan[7] , R. Höldrich[2] , and W. Aigner[1]

[1]Institute of Creative\Media/Technologies (IC\M/T), St. Pölten University of Applied Sciences, St. Pölten, Austria
[2]Institute of Electronic Music and Acoustics (IEM), University of Music and Performing Arts Graz, Graz, Austria
[3]Division for Media and Information Technology (MIT), Linköping University, Linköping, Sweden
[4]School of Creative Media, City University of Hong Kong, Kowloon, Hong Kong
[5]Ikerbasque Basque Foundation for Science, Bilbao, Spain
[6]Faculty of Engineering, Universidad de Deusto, Bilbao, Spain
[7]Scientific Computing and Imaging Institute (SCI), University of Utah, Salt Lake City, USA

Abstract

The research communities studying visualization and sonification for data display and analysis share exceptionally similar goals, essentially making data of any kind interpretable to humans. One community does so by using visual representations of data, and the other community employs auditory (non-speech) representations of data. While the two communities have a lot in common, they developed mostly in parallel over the course of the last few decades. With this STAR, we discuss a collection of work that bridges the borders of the two communities, hence a collection of work that aims to integrate the two techniques into one form of audiovisual display, which we argue to be "more than the sum of the two." We introduce and motivate a classification system applicable to such audiovisual displays and categorize a corpus of 57 academic publications that appeared between 2011 and 2023 in categories such as reading level, dataset type, or evaluation system, to mention a few. The corpus also enables a meta-analysis of the field, including regularly occurring design patterns such as type of visualization and sonification techniques, or the use of visual and auditory channels, showing an overall diverse field with different designs. An analysis of a co-author network of the field shows individual teams without many interconnections. The body of work covered in this STAR also relates to three adjacent topics: audiovisual monitoring, accessibility, and audiovisual data art. These three topics are discussed individually in addition to the systematically conducted part of this research. The findings of this report may be used by researchers from both fields to understand the potentials and challenges of such integrated designs while hopefully inspiring them to collaborate with experts from the respective other field.

1. Introduction

Over the course of the last few decades, two research communities have developed largely in parallel: one studying data visualization and one studying sonification. While the visualization community is primarily interested in "the use of computer-supported, interactive, visual representations of abstract data to amplify cognition" [CMS99], the sonification community studies "the use of non-speech audio to convey information" [KWB*99]. To this day, the communities seem to be largely disjoint, despite their shared goals promising fruitful collaboration. Both the theoretical cross-pollination [ERI*23, CLR22] and the practical integration and combination of sonification and visualization offer potential for interesting research outcomes. Therefore, with this state-of-the-art report (STAR), we want to shed light on audiovisual display idioms that sys-

tematically integrate data visualization and sonification. Informed by the definition of the visual idiom by Munzner [Mun15], we think of an *audiovisual display idiom* as "a distinct approach to creating and manipulating audiovisual representations of data."

In our daily lives, we perceive our surroundings in inherently multimodal and transmodal ways. We see, and we hear, we read books, and we listen to music. We use our senses to understand and explore the world around us. Although we are multisensorial beings, the predominant data analysis idioms are unimodal, often using visualization only. Sight and hearing are inherently different, with different strengths and challenges, and are most probably suitable for different approaches in data representation. Human visual perception can look upon a visual representation in a non-linear fashion, and the use of different types of graphs, charts, and other

K. Enge et al. / A State-of-the-Art Report on the Integration of Sonification and Visualization

visual formats can reveal patterns, correlations, and trends in data that are often not as noticeable in numerical form. Visual representations of data can also be experienced as more engaging and memorable to a user compared to tables with numbers [FPS*21]. Sonification, on the other hand, exploits the excellent ability of the human auditory system to recognize temporal changes and patterns. It is, therefore, useful when displaying complex patterns, such as changes in time and warnings for immediate action. In real-time environments, sonification allows a controller to perceive information without the constant monitoring of visual displays. Sonification also enables the communication of data and information for visually impaired individuals [WM10].

The human visual perception has some challenges that can be supported by sonification, and similarly, the auditory system has other challenges that, in turn, can be supported by visualization. Therefore, we believe that a well-designed audiovisual representation can be more than the mere sum of a visual and an auditory representation. To put this into practice, we refer to a typical real-world situation where the combination of visual and auditory inputs aids us in reaching a more informed conclusion: imagine rain falling outside of a window. It is often difficult to correctly estimate the density of rain by just looking outside a closed window. It is also not easy to estimate the amount of rain when only listening to it with your eyes closed. It is the holistic audiovisual perspective, perceptually integrating both of our senses, that allows us to best determine whether we should use an umbrella or even stay indoors.

Inspired by the capabilities of the human visual and auditory systems and the possibility of integrating visualization and sonification, this STAR covers academic contributions from both the visualization and the sonification communities that blend sonification and visualization within the context of data exploration and data presentation. We hope this STAR will help both visualization and sonification researchers realize the potential of such combinations and foster future collaborations between two often disjoint communities.

1.1. Sonification Background

This STAR being published at a visualization venue calls for a brief introduction to the sonification techniques that are part of our data (for a more comprehensive description of the different techniques, see the Sonification Handbook [HHN11]). Overall, we can distinguish the five main techniques of *audification*, *parameter mapping sonification*, *model-based sonification*, *auditory icons*, and *earcons*:

Audification is a technique to represent typically long sequences of data values (often time-series) by interpreting them as digital audio waveforms and directly playing them back over a loudspeaker. The resulting sound is a translation of the data values into the audible domain in terms of frequency and loudness. An example of audification is the playback of seismographic data with increased playback speed, such that the original low-frequency signal gets pitch-shifted into a range that is audible to humans. A design challenge for audification is the influence of the chosen playback speed on the salience of emerging auditory patterns that should inform a listener about their data.

Parameter mapping sonification (PMS) is a technique that in-

volves the association of data values with auditory parameters such as pitch, timbre, and loudness. The technique is conceptually closest to many standard visualization techniques as it employs the direct mapping of data values to auditory channels of a carrying sound. If a visualization utilizes a visual mark to represent information by the mark's position or its color, then a parameter mapping sonification utilizes a sound to carry information via its pitch, loudness, or other auditory channels. Similar to a visualization, such a sound is dependent on the mapping function between the data and the auditory channel as well as the nature of the mapping, for example, being linear to linear or linear to exponential. An open challenge in sonification research is the task-dependent and appropriate selection of auditory channels or the perceptual influence of one channel on another. In terms of visualization, this is comparable to the influence of the spatial size of a mark on our perception of its color.

The technique of *model-based sonification* (MBS) is inspired by the fact that most of the interactions with our physical environment result in an acoustic response. These acoustic responses, such as the sound of a drum being hit, inform us about the state of the object that we interact with (in the case of a drum, its tuning). Model-based sonification is a general term for sonification techniques that make use of dynamic models describing changes in a system over time. These dynamic models are "tuned" by the data that an analyst wants to explore. To listen to their data, a user is required to excite the model with an interaction, such as when a drummer needs to hit their instrument to hear its tuning. Another example is the excitation of a mass-spring model, where the mass and spring parameters are determined by high-dimensional data, defining the sound of the model when excited by a user. The technique is meant to foster exploratory data analysis, as the type and place of interaction strongly influence the acoustic response of the model. In this context, an ongoing research challenge is how to establish an intuitive relation between excitation modes (where and how hard to hit the drum) and typical interactions during exploratory data analysis, such as zooming or filtering of the data.

Also inspired by real-world sounds, *auditory icons* are short, distinctive sounds present in everyday life that can be compared to visual icons. This means that there is an inherent association between the auditory icon and the event they represent. A classic example of an auditory icon is the sound of a piece of paper being crumpled and thrown into a bin, representing deleting a file on a computer. A research challenge is the possibility of cultural differences between listeners, as they can lead to confusion or misinterpretation of auditory icons [JLSP15], similar to the way visual icons can be context-dependent.

Earcons, comparable to visual symbols, are short, distinctive sounds or melodies that are often used to represent specific events. These sounds are usually synthesized tones or sound patterns and can be described as designed or composed sound symbols. Since there is no inherent association with the real world, the meaning of the earcon needs to be learned before it can be beneficially used. Examples of such earcons are the sounds provided to pilots in the cockpit of an airplane, alerting them about events that require attention.

While other techniques for the sonification of data exist, these five are the most prominent ones. Within the scope of our STAR,

we identified all five of these techniques, even though the vast majority of papers utilize parameter mapping sonification. Some of the papers also combine techniques (for example, parameter mapping sonification and auditory icons, see Table 1).

A brief history of sonification: While the discipline of visualization has a relatively long history [Fri08], the research field of sonification is younger [Fry05]. In 1982, Bly submitted her PhD thesis Sound and computer information presentation [Bly82], where she suggested methods of encoding information into sound. Ten years later, the first International Conference on Auditory Display convened, which is regarded as the birth of the International Community For Auditory Display (ICAD). When its proceedings were published in 1994, the book Auditory Display [Kra94] was a reflection on the potential of the newborn field of research that encompasses sonification. Early on, Barrass [Bar97] presented a task taxonomy for auditory displays, called *TaDa!*, which stands for Tasks and Data. The *TaDa!* taxonomy is especially relevant in the context of this STAR, as it is well-aligned with taxonomies from the visualization literature [Ber83, BM13, SNHS13, Shn96, YKSJ07], and it also functions as inspiration for the classification applied later in this STAR.

The book Ecological Psychoacoustics, edited by Neuhoff in 2004 [Neu04], challenged many psychoacoustical studies (which is the part of psychophysics that involves the scientific study of sound perception, traditionally conducted in controlled laboratory environments). Neuhoff promoted an ecological sound approach to sonification from a holistic perspective, which echoes the aims of the BELIV workshop series established in 2006 at the Advanced Visual Interfaces conference [BPS06]. Neuhoff's intervention underscores the need to consider real-world contexts for transforming design principles and methodologies for auditory displays from theory to practice. This perspective emphasized integrating ecological factors in sonification to increase effectiveness and deepen the connection between auditory stimuli and real-world experiences.

The *sonification design space map* was introduced in 2007 by deCampo [dC07], guiding a designer's decision-making process of selecting an appropriate sonification technique for their task. The map creates a two-dimensional space between the number of data properties a designer intends to sonify and the number of data points that are necessary for the different sonification techniques to be employed adequately. Retrospectively, another milestone within the sonification community was the introduction of a now widely accepted definition of sonification as a scientific technique for representing data, presented by Hermann [Her08] in 2008. Before the introduction of this definition, it was less clear where to draw the border between artistic and scientific mappings from data to sound (which brings to mind the discussion that data visualization is more than just pretty pictures). As it reflects our understanding of the term sonification, we want to refer to the full definition below:

"A technique that uses data as input, and generates sound signals (eventually in response to optional additional excitation or triggering) may be called sonification, if and only if

- *The sound reflects objective properties or relations in the input data.*
- *The transformation is systematic. This means that there is a precise definition provided of how the data (and optional interactions) cause the sound to change.*

- *The sonification is reproducible: given the same data and identical interactions (or triggers) the resulting sound has to be structurally identical.*
- *The system can intentionally be used with different data, and also be used in repetition with the same data."*

The Sonification Handbook [HHN11], published in 2011, provided the first general and overarching perspective on the field of sonification, discussing both the theory and practice of sonification. The Handbook is still the most comprehensive collection of sonification work, and therefore, its publication year in 2011 also marks the beginning of the time period considered in our STAR.

Over the years, various design frameworks for sonification have been proposed. The design framework proposed by Barrass [Bar12] emphasizes the fusion of aesthetics and functionality to improve the accessibility and meaningfulness of sonifications for a broader audience. The work of Worrall in 2019 [Wor19b] formalized sonification techniques into a framework that also highlighted the challenges and advantages of these sonification techniques, as well as the importance of understanding processes and choices that influence sound representation. The *sonification design canvas*, introduced by Lenzi in 2021 [Len21], is a contribution to the construction of a more comprehensive design framework, with the aim of integrating all aspects into a cohesive design tool. Despite these efforts, developing a comprehensive protocol that systematically considers end-users at each stage of the design process has yet to be achieved.

More than 30 years after the beginning of systematic sonification research, we saw a considerable number of theoretical contributions to the field [Kra94, KWB*99, VH06, dC07, NW08, HHN11, GH12, Sup12, Nee19, Neu19, Wor19b, Len21]. Explicit work integrating sonification and visualization theory is rare, but has been called for [RW10]. In an attempt to find a common language and, consequently, build a theoretical bridge between the visualization and the sonification communities, Enge et al. [ERI*23] introduced three theoretical constructs to formally describe audiovisual display idioms. They defined the "auditory mark" inspired by the visual mark, the "auditory channel" inspired by the visual channel, and the "substrate of sonification" inspired by the spatial substrate that is used in visualization theory [Ber83, CMS99]. The definition of *time* as the substrate of sonification allows a description of sonification designs with auditory marks placed in time, with data encoded into their auditory appearance using auditory channels such as pitch or loudness. These definitions allow a high-level discussion and categorization of both the visual and the auditory part of an audiovisual display idiom. The theoretical constructs proved useful, so we adopted the term "auditory channels" for classification in this STAR as well.

1.2. Motivation and related work

Seminal visualization texts such as Wilkinson's Grammar of Graphics [Wil05] and Spence's textbook [Spe07] made clear statements that their understanding of visualization respectively graphics is not limited to vision but that data can be encoded for other sensory modalities such as sound. Already at the 6th IEEE Visualization Conference, Minghim and Forrest [MF95] postulated areas where sonification can help tackle visualization challenges such as adding complementary or redundant dimensions, natural mapping for time-oriented data, or improved memory of data. Some research agendas

[TC05] suggest multimodal approaches so that one sensory modality can overcome problems that others may have. In 1990, Grinstein and colleagues [GS90,SBG90] presented an audiovisual interface for the exploration of multivariate data using icons and parameter mapping sonification. Also, several works presented at early visualization conferences have integrated sound into visualizations for surfaces, volumes, and fluid dynamics [MF95,LWS96,RN96,VG97,FBZ*99]. The *Data Mountain* interface [RCL*98] augmented its spatial document management space with auditory cues that indicated how many pages were moving. Published as early as 1989, Gaver presented the *Sonic Finder,* which was the auditory user interface used in Apple computers and coined the term "auditory icon."

The auditory perception has an exceptional ability to detect temporal changes and patterns [GGB05,RG71]. Also, human hearing is capable of perceiving and distinguishing between several sounds simultaneously, at least between three auditory streams at the same time [SW13]. Another capability of our auditory perception is the possibility to detect and focus on events that spatially occur all around the listener, which can enable an information display to convey peripheral information to a user. Furthermore, the auditory system enables quick reactions when performing certain types of tasks due to the different processing times of the senses [JBKS15].

The combination of audio and visual data representations has shown to be advantageous, particularly for a number of applications. First, situations where the visual modality is busy with another task, such as monitoring [Ibe20, NB02], lend themselves beneficial for being complemented with sonification. Second, the combination of visual and auditory techniques has been shown to better facilitate learning [May14, SKS06], since it has the potential to increase working memory capacity and retention of information while also reducing cognitive load. Third, using sonification also shows benefits for data exploration. Flowers et al. [FBT97] published a seminal article in the sonification community that demonstrated that auditory scatter plots, where data is mapped to onset time and pitch of sounds instead of horizontal and vertical position of visual marks, can offer similar performance as visual representations.

Research has shown that visual and auditory perception is naturally integrated with each other, which can be observed, for example, with so-called crossmodal correspondences [Spe11, SDS23]. In our context, crossmodal correspondences describe a phenomenon where we perceive different visual and auditory stimuli as inherently related to each other. Such correspondences could provide an opportunity to use the strengths of the auditory modality to support and enhance visualizations, creating more effective and compelling representations of data [KBBG07]. Rosli and Cabrera [RC15], for example, have identified the potential of integrated designs to form a more concise, general representation of the data set compared to individual stimuli alone. Rubab et al. recently explored relationships between auditory channels and visual channels and suggested factors that impact their effectiveness [RYTW23].

Overall, the research described above suggests that the integration of the auditory modality has the potential to remedy challenges that exist for visual perception. Challenges, such as simultaneous brightness contrast [War19] or the Mach band phenomenon [LWP99], impact perception of visual representations [SGS*18, ZTSS23]. It has been demonstrated that various auditory channels can be success-

fully linked and related to visual channels [FB18, WHT06, CH04], which could offer a way of substituting visual channels with auditory channels to address these challenges.

1.3. Related Surveys

In general, systematic state-of-the-art reports are less established in the sonification community. A rare exception is the "*systematic review of mapping strategies for the sonification of physical quantities*" by Dubus and Bresin [DB13], which, however, just covers sonification-only contributions. Much earlier, in 2001, Walker and Lane [WL01] provided a website enabling researchers to search for sonification mappings that have been used in scientifically evaluated designs. Unfortunately, the website is no longer available. Another more timely and exhaustive exploration of sonification literature emerges in Andreopoulou and Goudarzi's 2021 publication [AG21]. The authors reviewed 456 papers from the International Conference on Auditory Display proceedings. This incisive analysis exposes compelling trends, ranging from the sonification domains to the diverse publication venues. The report reveals linguistic trends and explores the balance between research and artistic contributions. In addition, it illuminates the landscape of tools, methodologies, and evaluation practices that have led to sonification's multifaceted evolution. Marking the beginning of an important sociocultural reflection within the sonification community, in 2017, Andreopoulou and Goudarzi [AG17] also studied the "*representation of female researchers and artists in the conferences of the International Community for Auditory Display (ICAD)*". Their findings showed that only about 18% of ICAD papers were co-authored by women, with stagnant numbers between the years 1994 and 2016.

With respect to combinations of visualization and sonification, Caiola et al. [CLR22] recently presented an analysis of visual and auditory channels commonly used in audiovisual display idioms. Their survey includes combined idioms that map data attributes redundantly to both a visual channel (such as position) and an auditory channel (such as pitch). Analyzed works stem from the Sonification Archive (described below) and a Google keyword search using sonification-related terms exclusively. The *Sonification Archive*, widely known in the sonification community, is a curated collection of sonification designs, often related to other modes of representation, such as visualization. The Sonification Archive holds both artistic and academic contributions, as well as designs from data journalism.

Searching through the visualization literature, we were not able to find any STAR or survey focused on the integration of sonification and visualization. We explored the survey of surveys [ML17] but could not identify any related contributions. Therefore, to the best of our knowledge, this is the first systematic STAR dedicated to academic contributions in the intersection of sonification and visualization for data exploration and presentation.

1.4. How to Use This Survey

With this STAR, we intend to provide an overview of an emerging research field, as well as connect two mostly disjoint research communities. We hope to reach researchers from both communities,

inspiring them to intertwine sonification and visualization in their future research. We see several ways of using this STAR:

- using it as an overview, intended for researchers who seek a summary of the field.
- using it to find research opportunities and existing gaps in the field.
- using our supplemental material to study the existing meta-data in more detail, such as identifying authors from the respective other field for potential collaboration. Furthermore, we provide a public *Zotero* library, holding all relevant publication metadata, our tags, and all open access PDFs.

This STAR will be structured as follows: section 2 describes the methodology used to search and filter the literature identified as potentially relevant. In section 3, we describe our classification system and use it to discuss the survey literature. In section 4, we apply a meta-perspective on the survey data, describing correlations between individual tags, as well as the co-author network of the field. In section 5 we introduce the three adjacent topics of accessibility, monitoring, and arts, which are related to our STAR, but were not studied systematically. Finally, in section 6, we offer a concluding discussion focusing on future work.

2. Method

In this section, we discuss our inclusion and exclusion criteria and the methods we used to search for the relevant literature. We used a five-stage pipeline to construct a corpus of research that is at the intersection of visualization and sonification for data exploration and presentation (see Figure 1).

2.1. Scope of the Surveyed Literature

Sonification and visualization share the aim of making data interpretable to their users and observers. With this shared goal, combinations of the two can be designed for numerous possible applications and contexts. Our research interest in this STAR is the combination of sonification and visualization in the context of data analysis, covering both data exploration and presentation. A work that is relevant to our STAR must include both visualization and sonification of data. Therefore, a sonification with a visual interface that does not represent data is not enough to be considered relevant, and neither is a visualization with sounds that do not represent data. The work must be an academic paper published between the years 2011 and 2023 and must be peer-reviewed to be considered in our STAR.

Thinking more broadly about the combination of sonification and visualization, three additional areas of application come to mind: (1) accessibility, (2) real-time monitoring, and (3) arts. All three areas are vast, and a detailed classification of them is beyond the scope of our STAR. However, we find them relevant and inspirational for our field. Thus, we provide a brief introduction to the fields of accessibility, monitoring, and artistic contributions in section 5. In the context of accessibility, sonification can be used to support the collaboration between blind and sighted users by mapping data to both an auditory and a visual display. In the same manner, such a design could support the collaboration between deaf individuals and individuals without hearing impairment. Nevertheless, our research interest is the combination of sonification and visualization

for the integrated analysis of data. Therefore, in our STAR, we consider only designs intended to be used with both the visual and the auditory senses fully available to a user. The application of real-time monitoring, such as medical monitoring, critical infrastructure monitoring, alarms, or real-time feedback on body movement, is vast and distinct from the purpose of data exploration. Especially with respect to sonification and auditory display, the field is well-researched (e.g., [KIK19, SJMT19, VRGM20, WMY*17, HRM15]), and we will not cover such designs in this STAR. Artistic contributions have the potential to be highly inspirational for our field but require a different search method and, most likely, a different system of classification. Again, we decided not to systematically cover artistic contributions in this STAR but to provide a subsection discussing a list of representative works that may serve as a starting point for future research interests.

2.2. Search Strategy and Filtering

We base our corpus of literature on (1) publications that the authors have already been aware of from their previous work in this field and (2) an extensive online literature search. The online search was a keyword-based search in the digital libraries of IEEE Xplore, ACM Digital Library, and Springer Link, which include work published at IEEE VIS, CHI, and other VGTC- and SIGCHI-sponsored venues. Furthermore, we searched the digital libraries of Eurographics, ICAD, ISon, Organised Sound, and the Sound and Music Computing Community. Figure 1 provides an overview of the different stages we used to systematically filter our database for relevant publications. To keep track of the progress, we used a Google sheet document. The final database, including all papers and all tags by the authors, is provided in the supplemental material as a CSV file.

Stage 1 – The search query we used in stage 1 was the following: *("Visualization" OR "Visualisation" OR "Visual Analytics") AND ("Sonification" OR "Auditory Display") AND NOT (centrifug* OR lys* OR homogeniz*)*. It consists of an AND combination of a visualization term and a sonification term combined with the exclusion of specific word beginnings. While the terms "visualization" and "visual analytics" are well known, our audio-related search terms are likely to be less well-known overall. Nevertheless, we decided not to include search terms such as music, sound, or tone in our online search. Searching for publications with such broad terminology would have resulted in too many papers to be scanned. Also, we argue that academic publications interested in the integration of visualization and sonification are likely using the appropriate terminology. The reason to exclude papers that hold words starting with *centrifug*, *lys*, and *homogeniz* is the fact that sonification, also called sonication, is a term also used in biology to describe a process where sound is used to agitate particles in a sample. As papers using sonification in this context are not relevant to this STAR, we identified terms, including the three mentioned above, that are often associated with this meaning of sonification. The search in the Springer Link database was especially sensitive to the exclusion of these terms. As there could be false negatives using the exclusion of the three terms, we manually reviewed the paper titles that were excluded due to this strategy and restored five potentially relevant papers. This online search, combined with papers we were already aware of from our previous work in the field, resulted in a database

holding 1498 papers. We used the literature management software Zotero to download the respective papers and to make them available to all co-authors.

Stage 2 – In the second phase of our literature search, each paper title was read by two of the authors and classified into *potentially relevant* or *irrelevant*. We agreed to use an inclusive mindset for this stage, i.e., we tagged vague titles mostly as potentially relevant so as not to overlook many papers at this early stage. For papers that were tagged differently by two people, the two people came to an agreement, or the paper was taken to the next stage. Stage 2 resulted in a database holding 500 papers.

Stage 3 – In the third phase of our literature search, each abstract was read by two of the authors and classified into *potentially relevant* or *irrelevant*. Again, we used an inclusive mindset and tagged vague abstracts as potentially relevant. For papers that were tagged differently by two people, the two people came to an agreement, or the paper was taken to the next stage. Stage 3 resulted in a database holding 163 papers.

Stage 4 – In the fourth phase of our literature search, each paper was read by one of the authors and classified as *relevant* or *irrelevant*. As the authors had a solid common understanding of the inclusion and exclusion criteria by this stage, and as a person had the full paper information available to make a decision, one person decided on the relevance at this stage. The papers identified as relevant in stage 4 were classified using the tags explained in detail in Table 2. Each paper was first classified by one of the co-authors, and their tags were later verified by a second co-author. Whenever two co-authors initially disagreed on a specific tag, they came to an agreement, or the first author made a decision. Stage 4 resulted in a database holding 47 papers.

Stage 5 – Furthermore, we extended the corpus of relevant papers by snowballing [Woh14], checking all incoming and outgoing references of the articles matching our inclusion criteria. Snowballing was done by one of the co-authors with an exclusive mindset towards the paper titles, meaning that a vague title was not considered relevant. Stage 5 resulted in our database holding 57 papers overall, adding ten papers to the prior stage. During the final two stages, whenever we identified an audiovisual idiom that was published in more than one paper, we retained the most extensive version in our STAR. We identified two such cases where a design was previously published in a short paper but later expanded [YH18, MAFP19b]. It is notable that a considerable number of papers that are part of our final scope were published at the International Conference on Auditory Display (ICAD). This is not surprising as this venue is the most recognized venue to publish sonification work holding the largest single corpus of work in the field [BB12, AG21, GVEZ23]. It is also not surprising that it is a sonification venue that is prominent, as there is a structural imbalance between the domains. Many sonification designs, in general, also include some sort of visualization in their design, but that is not true the other way around.

3. Categorization and Results

In this section, we will discuss the relevant literature in detail from the perspectives of our classification. The systematic integration of sonification and visualization is a wide and diverse research field
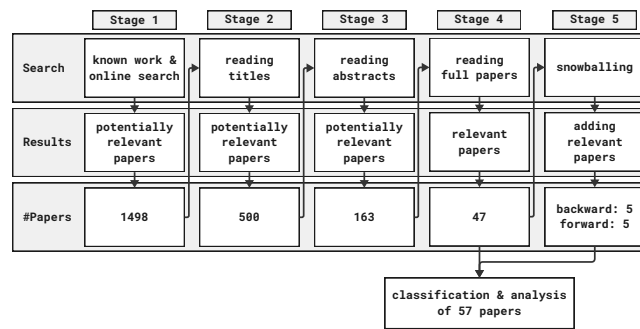
Figure 1: *The literature search was conducted in five stages, ranging from an online keyword search to using the snowballing technique on the papers identified as relevant. Figure inspired by [YM22].*

that is difficult to classify using only a handful of categories. Therefore, we decided to apply an extensive list of tags to the literature to be able to present diverse perspectives on the field, mostly concerning basic research, i.e., the basic principles that distinguish the designs/idioms from each other.

To give readers an initial thematic overview of the field, we will start by briefly describing each of the 57 selected papers in subsection 3.1. We will do so by clustering the literature to the following topics: astronomy, medicine and health, molecular science, earth science, domain agnostic data displays, and other topics. We will then continue with a discussion on the purpose that an idiom can be designed for (subsection 3.2), followed by an analysis of idiom design possibilities (subsection 3.3). We will study several technical perspectives that were tagged individually for the sonification and the visualization aspects of each paper: the reading levels (subsection 3.4) suggested by Bertin [Ber83], the search levels (subsection 3.5) suggested by Munzner [Mun15], as well as the dataset types [Mun15] including the levels of measurement of the displayed data (subsection 3.6). We will then review different levels of mapping redundancy (subsection 3.7), different evaluation methods (subsection 3.8), different target platforms, and various possibilities of interacting with audiovisual display idioms (subsection 3.9). Finally, we study the diverse user groups and the possible goals of designers (subsection 3.10).

All the above categories and subcategories are concisely presented in two tables. Table 1 shows an overview of all papers and a subset of their most relevant tags. Table 2 provides a detailed description for each category and their respective subcategories. We also report on the total number of papers within each subcategory. The four technical categories mentioned above have distinct visualization and sonification tags, which are represented in the left and right boxes in the "Num." column. The luminance of the boxes encodes the total number of papers in each subcategory.

3.1. Thematic Corpus Overview

Before we employ the classification system described in Table 2, we want to provide a thematic corpus overview. This overview is intended for readers who are interested in a special field of application, such as astronomy or earth sciences. While we do not intend to

Paper	Theme	Demo Link	Purpose			Redundancy			Visualization Idiom	Sonification Technique					Goal			
			Exploration	Presentation	Both	Redundant	Mixed	Complementary		PMS	MBS	Audification	Auditory Icons	Earcons	Education	Data Analysis	Research	Public Engagement
[EEBR21]	Astronomy	[<—>]			•	•			volume rendering	•								•
[Rib19]	Astronomy	[<—>]		•		•			line chart, scatter plot	•						•		
[HTHB22]	Astronomy	[<—>]		•		•			point cloud	•								•
[HPDW23]	Astronomy	[<—>]			•			•	scatter plot	•		•			•			
[Rib18]	Astronomy	[<—>]			•	•			volume rendering	•						•		
[RS22]	Astronomy	[<—>]		•				•	3D map	•								•
[TB23]	Astronomy			•				•	3D scatter plot	•								•
[BTB23]	Domain Agnostic Display		•			•			line chart, parallel coordinates	•						•		
[CWM21]	Domain Agnostic Display	[<—>]			•	•			line chart	•					•	•		•
[DLVDCG22]	Domain Agnostic Display	[<—>]			•	•			line chart	•						•		
[DCM*18]	Domain Agnostic Display				•	•			convex hull, other, bar chart	•						•		•
[ERI*22]	Domain Agnostic Display	[<—>]	•			•	•	•	scatter plot	•						•		
[FBC12]	Domain Agnostic Display				•	•			line chart	•						•		
[GKW21]	Domain Agnostic Display	[<—>]	•					•	network	•						•		•
[KLTW17]	Domain Agnostic Display	[<—>]		•			•		line chart	•								•
[LF21]	Domain Agnostic Display	[<—>]			•	•			line chart	•					•	•		•
[MAFP19b]	Domain Agnostic Display	[<—>]*	•			•			heatmap	•						•		
[PCB23]	Domain Agnostic Display	[<—>]			•	•			line chart	•					•			
[PC19]	Domain Agnostic Display			•		•			line chart	•						•		
[RJ16]	Domain Agnostic Display	[<—>]	•			•			scatter plot, parallel coordinates	•						•		
[YH18]	Domain Agnostic Display	[<—>]	•					•	scatter plot		•					•		
[FN18]	Domain Agnostic Display		•				•		line chart	•						•		
[Bal15]	Earth Science			•			•		heatmap	•								•
[Bea11]	Earth Science	[<—>]	•			•			map	•						•		
[BF12]	Earth Science	[<—>]	•			•			map	•						•		
[GDAS*18]	Earth Science				•	•			map	•							•	
[HK22]	Earth Science			•		•			geographic scatter plot	•								•
[HCTP14]	Earth Science			•		•			heatmap			•			•			
[MMU16]	Earth Science		•					•	map	•		•				•		
[NRL*12]	Earth Science	[<—>]	•					•	map	•					•			
[PFH*22]	Earth Science	[<—>]			•		•		dot map			•					•	
[SAR22]	Earth Science	[<—>]		•		•			fluid-like simulation	•								•
[WW15]	Earth Science	[<—>]		•		•			line chart			•						•
[GRK*16]	Medicine and Health		•					•	slicing, volume rendering	•						•		
[GR11]	Medicine and Health		•			•			slicing	•						•		
[LSB*23]	Medicine and Health	[<—>]	•				•		map	•						•		
[MNW*18]	Medicine and Health	[<—>]	•					•	volume rendering	•						•		
[RFM13]	Medicine and Health		•			•			slicing	•						•		
[TMN*21]	Medicine and Health	[<—>]	•					•	volume rendering	•						•		
[AJB*18]	Molecular Science	[<—>]	•					•	3D molecule rendering	•						•		
[BBV16]	Molecular Science		•					•	3D molecule rendering	•			•			•		
[BM20]	Molecular Science		•			•			volume rendering	•						•		
[RFK*15]	Molecular Science		•					•	volume rendering	•			•			•		
[LLW21]	Others	[<—>]	•					•	3D network	•								•
[NSC16]	Others	[<—>]	•				•		gantt chart	•				•		•		
[CB17]	Others			•		•			bar chart	•						•		
[MMM18]	Others	[<—>]		•		•			individual circular design	•								•
[ASH*12]	Others	[<—>]	•			•			volume rendering	•						•		
[Her20]	Others	[<—>]	•			•			network	•							•	
[PBM15]	Others		•					•	3D network	•			•			•		
[PBV14]	Others		•					•	network	•						•		
[HAR16]	Others		•				•		dotted chart visualization	•				•		•		
[JMP13]	Others				•			•	3D scatter plot	•						•		
[KAV21]	Others	[<—>]			•	•			3D point cloud	•								•
[Rön21]	Others	[<—>]		•				•	bar chart	•								•
[AK11]	Others			•			•		map	•						•		
[BB19]	Others		•			•			point grid	•						•		

Table 1: *A table showing all 57 entries in our database, sorted by their thematic cluster. In the columns for the used sonification technique, "PMS" stands for parameter mapping sonification, and "MBS" stands for model-based sonification. Where available, the demo links point to the supplemental material of the papers (last accessed on 22nd of December 2023). *There are two additional demo videos associated with this paper: [<—>] and [<—>]; all three demos require the password: icad2019.*

provide detailed descriptions of the individual articles at this point, we will present more detailed descriptions in the later subsections 3.2 to 3.10.

Scanning our database, we identified seven **astronomy** related papers. Riber [Rib18] presented a prototypical virtual and interactive audio synthesizer called *Planethesizer* that enables its users to design sonifications, especially focused on planetary data. *Sonifigrapher* [Rib19] is a virtual synthesizer that lets users sonify the light curves data from NASA's exoplanet archive. Also, the recently presented *Sonified Hertzsprung-Russel Diagram* [HPDW23] sonifies the light curves, with the diagram acting as both the visualization and the interface to choose a star to be sonified. With this design, hearing a constant pitch will inform a user about the rotation of a star. The rotation, temperature, and other parameters of planets in our solar system were also sonified by Elmquist et al. [EEBR21] in *OpenSpace Sonification*. Their design can be used both with conventional computer desktop environments and in planetarium settings, and they are tailored towards public outreach and science communication. Public outreach is also the core of the publication *Audio Universe* by Harrison et al. [HTHB22]. The publication describes the design of a 35-minute audiovisual show about the solar system integrating visualization and sonification, as well as an audiovisual animation displaying the stars in the same order they appear to our eyes during dusk. Similarly, Russo and Santaguida [RS22] collaborated with NASA, celebrating the discovery of the 5000th exoplanet. Their design displays the exoplanets as they were discovered over the years. Recently, Traver presented another audiovisual installation where users can control the auditory representation of the planets using a Midi controller [TB23].

We identified six **medicine and health** related topics in our database, out of which three are related to brain scans [GR11, RFM13, GRK*16], two are audiovisually displaying blow flow and aneurysm models [MNW*18, TMN*21], and one is concerned with Covid-19 data [LSB*23].

Our database holds four idioms that we relate to **molecular science**. In an idiom presented by Rau et al. [RFK*15], scientific visualization of a molecular simulation is enhanced using parameter mapping sonification and auditory icons. Among other things, their design guides the attention of a user towards visually occluded phenomena using sonification. Ballweg et al. [BBV16] use sonification with the intention of supporting chemists and structural biologists with drug design. For their sonification plug-in, they focused on tasks that were not well supported visually in a software for the interactive visualization of molecular structures called "UCSF Chimera." In the context of biomolecules simulation, Arbon et al. [AJB*18] developed a sonification displaying characteristics of the "free energy landscape," a map used to study the properties of biomolecular systems. Their technique allows a user to visually inspect the physical configuration of a biomolecule while listening to their corresponding free energy landscape. Exploring the possibilities of 3D sound, Bouchara and Mones [BM20] suggested a work-in-progress immersive sonification model to study protein surfaces.

The **earth science** cluster in our survey data holds eleven publications covering topics that range from oceanographic data [SAR22], wildfires [HK22], hurricanes [Bal15], and climate change [Bea11], to sonophenology [NRL*12], seismology [HCTP14, WW15, MMU16, PFH*22], and geospatial data displays [BF12, GDAS*18].

The category of **domain agnostic data display** idioms in our survey data holds 15 papers. These idioms are not designed to support users from a specific domain but are implementations tackling problems across multiple domains. Six of the papers describe software frameworks that are intended to help people design sonifications along with visual representations of their data [PC19, PCB23, LF21, DLVDCG22, CWM21, KLTW17]. Their unifying core goal is the democratization of sonification as a technique to represent data, hence making it accessible to more people, both professionals as well as domain experts. Other publications focus on basic research combining different sonification techniques such as parameter mapping sonification or model-based sonification with basic information visualizations such as scatterplots [ERI*22, RJ16, YH17], parallel coordinates plots [BTB23, RJ16], or line charts [FBC12, FN18] to study the potentials of audiovisual display idioms. Groppe et al. [GKW21] studied network visualization and sonification through their design, while Malikova et al. [MAFP19b] show the potential of sonification to help users identify smallest symmetry differences in scalar fields visualizations. While most studies focus on metrics such as precision, error rates, or task completion times, the study by Du et al. [DCM*18] explicitly investigates the sonification's influence on user engagement.

Finally, we want to provide a brief overview of the remaining **other** 14 publications not part of the above thematic clusters. The topics related to the publications tagged as "other" are diverse, ranging from a multimodal system for analyzing business process execution data [HAR16] or Git version control data [NSC16], to an audiovisual representation of the Portuguese consumption patterns [MMM18], to a multimodal implementation of "Game of Life" [KAV21], and a musical sonification aimed at conveying information about running data and emotion [Rön21]. Several publications focus on audiovisual data representation in virtual or extended realities [CB17, PBV14, PBM15, BB19, JMP13] and two publications focus on explainable AI [LLW21, Her20]. Alonso et al. [ASH*12] presented an interface for product design that communicates a virtual object's geometrical shape using visual, haptic, and auditory stimuli, and Adhitya and Kuuskankare [AK11] proposed a prototype that offers a sonification-based approach to urban design planning.

As we have now presented a thematic overview of all 57 papers that are part of our database, we can focus on the meta-level classification in the following subsections. With these more high-level descriptors, we intend to provide a number of versatile perspectives on the literature. They will help us identify research gaps and opportunities for the systematic integration of sonification and visualization for the future work of both research communities in section 6. Along the discussion of these perspectives, we will present selected papers that are representative examples for the respective category. By doing so, we intend to provide the reader with (1) a broad exploration of the field overall and (2) insights into the content of the actual papers themselves.

Category	Num. (Vis)	Num. (Son)	Subcategory	Explanation
Purpose	29		Exploration	designs used for data exploration
	17		Presentation	designs used for data presentation
	11		Both	designs used for both purposes above
Visualization Design			Idiom	the visualization idiom, such as scatter plot or line chart
			Identity channels	the name of the visual channel, such as color hue, or shape, that is used to communicate the identity of an item (i.e., "What" something is)
			Magnitude channels	the name of the visual channel such as position or length that is used to communicate the magnitude of an attribute (i.e., "How much" something is)
Sonification Design			Technique	the sonification technique such as parameter mapping sonification or earcons
			Identity channels	the auditory channel such as timbre that is used to communicate the identity of an item (i.e., "What" something is)
			Magnitude channel	the auditory channel such as pitch or duration that is used to communicate the magnitude of an attribute (i.e., "How much" something is)
Reading Level	50	29	Whole	designs that display *all* datapoints
	13	29	Group	designs that display a *group* of datapoints
	4	17	Single	designs that display *single* datapoints
Search Level	2	2	Lookup	the user knows the location and the target of the search (see Figure 4)
	9	6	Browse	the user knows the location but not the target of the search
	7	7	Locate	the user doesn't know the location but the target of the search
	40	42	Explore	the user doesn't know the location or the target of the search
	11	8	None	none of the above
Dataset Type	29	40	Table	data constructed from items and attributes (spreadsheets; see Figure 5)
	5	3	Network	data constructed from items (nodes), links and potentially attributes
	9	7	Field	data constructed from grids (positions) and attributes
	20	9	Geometry	data constructed from items and positions
Level of Measurement	22	13	Nominal	data that builds categories (such as different fruits)
	8	8	Ordinal	data that builds ordered categories (such as t-shirt sizes)
	27	31	Interval	data that has equal intervals, such as the time on the clock
	27	34	Ratio	data that has equal intervals and a meaningful zero point, such as length or weight
Level of Redundancy	17		Redundant	a design mapping all displayed information to both senses redundantly (see Figure 6)
	14		Complementary	a design mapping part of the information exclusively to the visualization and another part of the information exclusively to the sonification
	28		Mixed	a design mapping some information redundantly, some information complementary
Evaluation System	12		User Performance	evaluations collecting metrics such as error rates or task completion times
	11		User Experience	evaluations collecting user feedback, typically done in a usability test
	2		Algorithmic Performance	evaluations doing measurements without users such as rendering speed
	17		Qualitative Result Inspection	evaluations providing subjective arguments on the quality of the result
	20		None	no evaluation
Target Platform	43		Desktop Computer Display	conventional screen on desktop computer
	4		XR	extended reality settings such as virtual or augmented reality glasses
	10		Physical Environments	environments that foster collaboration such as a CAVE
	5		Touch Display	interactive screen that users can interact with via touch
User Interaction	37		Yes	designs that require user interaction other than pressing "play"
	20		No	designs that require no user interaction
Users	25		Domain Experts	domain experts that are not visualization or sonification researchers
	7		Researchers	visualization or sonification researchers
	31		General Public	the interested general public
Goals	8		Education	idioms designed for education
	35		Data Analysis	idioms designed for data analysis (incl. data exploration)
	3		Research	idioms designed for visualization and sonification research
	17		Public Engagement	idioms designed for public engagement
Demo	29		Yes	the paper links to a demo such as an interactive website, a video, or an audio recording
	23		No	the paper doesn't link to a demo
	5		Yes, but not online	the paper provides a link that is not online anymore

Table 2: *Descriptions and distributions of all used classification tags. For the categories with two separate tags under the "Num." column, the left and right boxes represent the visualization and sonification distributions, respectively.*

3.2. Purpose

In this section, we will discuss two different purposes an audiovisual idiom can be designed for, inspired by the taxonomy described in [Mun15]: exploration and presentation. The *purpose* should be read as the general design goal of a tool with a broad perspective. Therefore, the term exploration also covers what is widely referred to as data analysis.

3.2.1. Exploration

A wide range of audiovisual idioms in our database support users with the exploration of data (see Table 2). We refer to the purpose of exploration in cases where a user intends to acquire new knowledge from their data by using an audiovisual display idiom.

Exploratory data analysis is the core aim of model-based sonification [Her02]. The *Mode Explorer* is an interactive audiovisual display idiom, combining a scatterplot visualization and model-based sonification to explore high-dimensional data [YH18]. To explore the data, a user can "scratch" a scatter plot of dimensionality-reduced data with an interactive pen. The scratching introduces a virtual particle to the high-dimensional space, which will follow a gradient descent to the nearest mode in the data. A user will be able to hear the "kinetic energy" of this virtual particle. While the particle travels closer and closer to its final mode, the sound gradually turns more and more harmonic, finally resulting in a clear pitch when a mode is reached.

AIive combines visualization and sonification in a virtual reality environment [LLW21]. The idiom helps users understand the basic concepts of neural networks. Users can manipulate the weights of the nodes of a virtually displayed 3D neural network by dragging the nodes in 3D space. In real-time, the sonification displays the loss and accuracy of the neural network, therefore enabling a user to explore different constellations of node weights. Users are also able to add or delete nodes, hence experimenting with the complexity of the neural network. While the authors do not provide a user study in their paper, it is most plausible the exploratory character of the design can support people in understanding the basics of neural networks.

In a recent publication, Lemmon et al. presented an audiovisual map idiom that seeks to tackle some of the sociotechnical challenges associated with epidemiological mapping [LSB*23]. Using their idiom, users can interactively explore Suffolk County's experience with Covid-19. The black population and associated case numbers are displayed on the left audio channel, while the white population and associated case numbers are displayed on the right channel via the pitch of sine and triangle waves. A correlation between black and white population numbers and COVID-19 case numbers and their dependency on different regions becomes clearly apparent when listening to the sonification while brushing the map using the computer mouse.

3.2.2. Presentation

We refer to the purpose of presentation when a user intends to present prior knowledge to others or in cases where a user is presented with information that is new to them but not to the designer.

One potential application of a presentation-only approach is in citizen science, exemplified by the communication of significant discoveries, such as NASA's announcement of the 5000th exoplanet [RS22]. This representation is explicitly crafted for communication with the general public, primarily through social media channels. Translating data into various sensory modalities is straightforward, while significant attention is devoted to achieving a pleasing and harmonious aesthetic. The sonification, complemented by two animated videos, enhances visual comprehension, vividly portraying celestial spheres with dynamic elements representing exoplanet characteristics. NASA's strategic dissemination on social media platforms garnered substantial engagement, with positive feedback indicating the presentation's broad appeal.

The presentation objective appears linked to eliciting emotional engagement from end users, as evident in Rönnberg's research [Rön21]. This representation involves assessing running statistics, weather data, and associated emotions. Presented as an animated visualization synchronized with sonification, it depicts weekly runs emphasizing evoking emotional responses rather than precise data interpretation—a practice denoted by the author as "musification" (see further discussion about sonification and music in Vickers, 2016 [Vic16]), leveraging sound's dual capacity to illustrate data and engender a musical experience. The study assesses the representation's efficacy through a user study, wherein participants watched a video. Results indicate that the sonification effectively conveys intended emotions but at the expense of a less accurate data representation.

Recent works delve into environmental concerns and the promotion of awareness. One such noteworthy contribution is presented in [HK22]. This publication centers around a multimodal museum installation designed to foster public engagement with wildfire forecasts in specifically chosen California and South Korea regions. The project incorporates interactive data visualization, sonification, and 3D-printed sculptures. Combining these elements conveys wildfire data, creating a comprehensive and immersive experience. The installation allows audiences to explore the representation freely through contactless interaction. Without user engagement, the wildfire data representation seamlessly loops across the screen. Showcased at the ARKO Art Center in Seoul, this project successfully captivated the attention of 20 users. The outcomes revealed a heightened level of interest and comprehension regarding the impact of wildfires through the effective utilization of multimodal interaction. The project emphasizes observations and prompts audience reflection, leveraging new media to enhance public climate change comprehension.

3.2.3. Exploration and Presentation

Some of the designs we identified were designed both to explore and to present data. Inclusion in this category requires the designer to put similar weight on the exploration and presentation of the data. To present a few examples, Paté et al. [PFH*22] demonstrate their audiovisual display idiom with three seismic data sets, where the sonification methods are adapted to the specific properties of each data set. A multi-scale audification method is presented where different speed factors are used depending on the size of each dataset, and different sound designs are used to highlight specific properties of the data. Huppenkothen et al. [HPDW23] created an audiovisual

version of the Hertzsprung-Russell Diagram where the user can listen and compare auditory representations of each type of star that is included in the diagram (see Figure 7(c)). It is also possible to filter the diagram according to certain criteria, which, in turn, filters which sonifications can be listened to. Elmquist et al. [EEBR21] created complementing sonifications of the planets in the solar system, which were integrated with a visualization, where a user can explore the properties of the planets and make comparisons between them. The user can listen to all of the sonifications for each planet at the same time or enable specific sonifications for each planet to compare specific properties. Du et al. [DCM*18] conducted a study investigating the enhanced visualization of basketball player movement data during a game. The visualization is designed to convey the offensive and defensive dynamics of a team. The interface primarily enables users to specify a particular time range of interest, providing a more comprehensive view of information within that timeframe. Notably, sonification is integrated exclusively during the exploration phase.

3.3. Audiovisual Idiom Design

Designers of audiovisual idioms have countless possibilities to combine their visualizations and sonifications. They can choose from a vast number of established visualization idioms and sonification techniques. Having chosen two designs to integrate with each other, they are free to choose the visual and auditory channels they want to map their data to. During the classification of the surveyed literature, we used six different categories to capture the state-of-the-art of audiovisual idiom design. The categories are the visualization idiom, the sonification technique, the visual identity- and magnitude channels, as well as the auditory identity- and magnitude channels. The constructs of identity- and magnitude channels are used to distinguish between encodings that communicate "what" something is and "how much of" something there is [Mun15, ERI*23]. Typical visual identity channels are the shape of a visual mark or its color hue. Typical visual magnitude channels are the position or the length of a visual mark. A typical auditory identity channel is the timbre of a sound, often generated using different musical instruments. Typical auditory magnitude channels are pitch and loudness [DB13, CLR22]. In the following, we will make use of these descriptions to shed light on some of the designs in our database.

SoniScope [ERI*22] is an interactive audiovisual display idiom that combines a scatterplot visualization with an interactive parameter mapping sonification (compare Figure 2). Inspired by the technique of auscultation, the SoniScope acts as a stethoscope for data, providing an auditory lens to "listen into" one's data. The exploratory character of the idiom is most apparent when the sonification displays non-visual data dimensions. To interact with the idiom, a user can brush the data using a visual lens. Clicking into the scatterplot will then trigger the sonification of a non-visual attribute of the selected data items. Regarding the used magnitude channels, we see a common combination of visual position and the pitch of the sounds, also employed in publications such as [BF12, DLVDCG22, FBC12, FN18, KLTW17].

Bearman [Bea11] studied the possibilities of displaying the uncertainty in future climate projections for the UK. The proposed audiovisual idiom shows a map of the UK with a heat map overlay
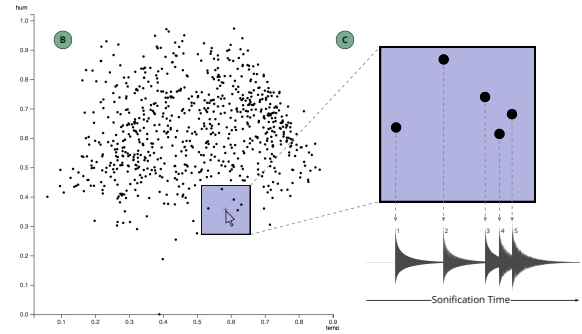
Figure 2: *The SoniScope audiovisual display idiom [ERI*22]. Users can interactively select a region in a scatterplot to sonify an additional data attribute for the respective items.*

displaying the projected temperature values. This visualization is combined with the sonification technique of interactive parameter mapping. People in a user study were asked to hover over the heat map with a mouse, triggering the sonification of the respective region. A higher pitch of a trumpet sound communicated a higher uncertainty for that region. Hence, for their design, Bearman used the spatial position as a visual identity channel and color as a visual magnitude channel. This sonification did not use an identity channel (only one sound could be heard) but used pitch as a magnitude channel.

Several papers in our database describe frameworks that are explicitly developed for the design of sonifications (see examples in Figure 3), always in combination with a visualization [PC19, PCB23, LF21, DLVDCG22, Rib19, CWM21]. As a representative and flexible example we want to discuss the *Highcharts Sonification Studio* [CWM21], a collaboration between the company Highcharts and the Georgia Tech Sonification Lab. Regarding the design of the visualization, the environment offers line- and area charts, as well as scatter plots, bar charts, and pie charts, all with their respective standard visual channels such as position, length, angle, or color hue. The sonification is done using parameter mapping with several options for auditory magnitude channels such as pitch, spatial position in the stereo field, loudness, or harmonic range. Different attributes in the data can be distinguished using different musical instruments, hence the employed auditory identity channel is the timbre of the different sounds. The default option of the Highcharts sonification studio is set to combine a line chart with an auditory graph, which traditionally plays back the line from left to right using pitch over time.

In their paper, Winters et al. [WW15] describe the design of the visualization and the sonification of the 2011 Tohoku earthquake in Japan. The design combines four line charts of four seismographic recording stations in Japan with an audification of the same data streams. The essence of the audification is that the low-frequency recordings of the earthquake are played back at a faster tempo, making them audible to the human ear. This very direct mapping between the physical (non-audible) phenomenon and the audification to the audible range results in a rich auditory impression that becomes informative in an ecologically valid manner. The paper

(a)   (b)   (c)
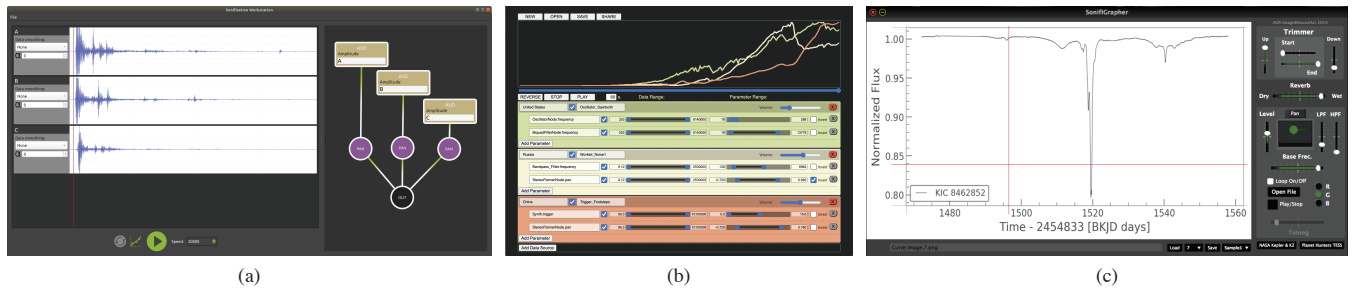
Figure 3: *Examples of frameworks for sonification in combination with visualization. (a) The Sonification Workstation by Phillips and Cabrera [PC19] is designed for general sonification tasks. (b) The WebAudioXML Sonification Toolkit (WAST) user interface by Lindetorp and Falkenberg [LF21] also aimed for general sonification tasks. (c) The Sonifigrapher virtual synthesizer interface by Riber [Rib19] for sonifying light curves. All screenshots CC BY 4.0.*

mostly discusses the popularity of this sonification on YouTube, currently with around 90k views, explicitly mentioning the relevance of combining the sonification with visualization for its success in public outreach. The visualization uses position as both its identity and its magnitude channel. The sonification employs the spatial stereo position of the four channels to distinguish between them, hence as their identity channel. The magnitude channels are a mix of pitch and timbre, resulting from the direct mapping between the physical phenomenon and the sound.

Discussion: From the data, we see that parameter mapping is by far the most used sonification technique for the design of audio-visual idioms – 53 out of 57 papers in the database employ some sort of parameter mapping sonification. Only seven of them combine parameter mapping sonification with other techniques such as audification [MMU16, HPDW23], earcons [HAR16, NSC16], or auditory icons [BBV16, RFK*15, PBM15]. The technique of audification is part of the database three times on its own [PFH*22, HCTP14, WW15], and once in combination with earcons in a second design of a paper [PFH*22]. Yang and Hermann are the only authors employing model-based sonification with their *Mode Explorer* design [YH17]. The most prominent auditory identity channel in our literature corpus is timbre (used 21 times), and the most used auditory magnitude channel is pitch (47 times), also used in combination with other channels. These findings are just in line with other meta-analyses of the field of sonification [CLR22, DB13]. The visualizations in our corpus employ idioms such as line charts (10 times), scatter plots (8 times), and maps (7 times), as well as networks, bar charts, heatmaps, and several other idioms. These publications use the identity channels color hue (27 times), position (11 times), as well as shape (11 times). The employed visual magnitude channels are position (33 times), color hue (15 times), size (9 times), and other channels such as opacity, tilt, or animation.

Tagging the surveyed literature made us realize that using the concept of the "channel" for techniques such as audification or model-based sonification is not trivial. With these techniques, the designer of a sonification, to some extent, loses control over its sonic outcome. This contradicts the definition of a visual channel as "a way to control the appearance of marks" [Mun15]. Instead, thinking of the "channel" as "the quality of a mark that transports information," also allows for the description of audifications and

Sonification	Visualization		
	whole	group	single
whole	27	3	1
group	24	12	1
single	14	4	3

Table 3: *The number of entries for Bertin's different reading levels [Ber83] and their distribution to the two display types.*

model-based sonifications. While these techniques typically result in highly complex sound sequences, it will often be qualities such as pitch or timbre that are informative to the listener. In general, we see a quite diverse field of different audiovisual idiom designs in our corpus.

3.4. Reading Levels

The classification of designs with respect to reading levels is inspired by the taxonomy by Bertin in his seminal book Semiology of Graphics [Ber83]. The three reading levels describe the amount of data a user studies using a specific tool. The "Whole" level describes tools that enable the user to ask questions about "all" of the data under consideration. The "Group" level describes tools that enable the user to ask questions about a "subgroup" of the data under consideration, and the "Single" level describes tools that enable the user to study "single" items. For this category, we decided to assign two tags per entry in our database: One for the visualization part and one for the sonification part of the idiom. The classification using reading levels aims towards potential differences in the distribution of tasks between sonification and visualization, such as overview and detail phenomena.

Rau et al. [RFK*15] presented an audiovisual idiom that lets users interactively explore molecular structures by using a "virtual microphone" that can be placed inside a 3D molecular visualization. The visual design provides an overview using the "whole" level while displaying details at the "group" level via the sonification. At the same time, the design provides a user with information about potentially visually occluded data. Hence, it makes the user aware of the existence of a phenomenon that they could, if relevant, study

Figure 4: *The four search levels suggested by Munzner: lookup, browse, locate, and explore. (Figure from "Visualization Analysis and Design" [Mun15] by Tamara Munzner, with illustrations by Eamonn Maguire, CC BY 4.0.)*

Sonification	Visualization			
	explore	browse	locate	lookup
explore	37	4	5	5
browse	3	2	2	2
locate	4	3	5	5
lookup	1	0	1	2

Table 4: *The number of entries for Munzner's different search levels [Mun15] and their distribution to the two display types.*
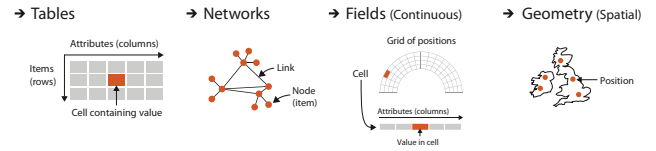

Figure 5: *Four dataset types suggested by Munzner: tables, networks, fields, and geometries. (Figure from "Visualization Analysis and Design" [Mun15] by Tamara Munzner, with illustrations by Eamonn Maguire, CC BY 4.0.)*

Sonification	Visualization			
	table	network	field	geometry
table	23	1	4	12
network	0	2	0	0
field	0	0	7	1
geometry	1	0	0	9

Table 5: *The number of entries for the different dataset types [Mun15] and their combinations between the two display techniques. Notably, the visualization of geometries is often combined with the sonification of table data.*

in detail at a later point. The metaphor of the microphone to be used to "listen into the data" is similar to the one of the stethoscope in [ERI*22], where the visualization provides an overview while users can interactively choose a subset of data to display acoustically. An example of an idiom working at the sonification "whole" level and the visualization "group" level is the *Mode Explorer* [YH17]. The design sonifies a high-dimensional data space while visualizing a two-dimensional projection of that space with a scatterplot.

Discussion: Out of all possible combinations of reading levels and display type, the majority of cases (27) use the "Whole" level for both the sonification and the visualization (some of them are [EEBR21, MMM18, RS22, SAR22, WW15]). The next biggest group in our database holds 24 papers that can display data at the "Whole" level using their visualization and at the "group" level using their sonification, such as [BF12, BTB23, FBC12, GDAS*18, RFK*15]. Generally speaking, we can observe that the visualization parts are mostly covering the same or a higher level of reading than the sonification parts of a design (compare Table 3). This phenomenon can be directly related to the distribution of tasks on the two display types, with visualization more often providing an overview and the sonification rather displaying details. In section 6, we will argue for breaking such patterns in the future work of the community.

3.5. Search Level

Inspired by the taxonomy suggested by Munzner [Mun15], we tagged four different search levels, again assigned individually for the sonification and the visualization parts of the idiom. The type of search a user applies depends on their prior knowledge (see Figure 4). Users who know what they are looking for and where they can find it will do a *lookup*. Searching for an unknown pattern at a known location is called *browsing* while *locating* is a search

without knowing the location but the pattern one is looking for. A search uninformed both with regards to the location and the type of pattern one is looking for is called *exploration*. Designs that combine sonification and visualization could serve each of the described search levels as well as a combination of them.

Malikova et al. [MAFP19b], for example, presented an idiom that helps users explore scalar fields. The regional magnitude of small areas of a scalar field is sonified using pitch. From the search level perspective, a user of this idiom does not generally know what to look or listen for at the beginning of the analysis, and therefore, both modalities are used to explore the presented data. The system could, for example, reveal the existence of small symmetry differences in scalar fields. What visually seems to be a perfectly symmetrical field could become apparent as not quite symmetric when two similar pitches result in an acoustic phenomenon called "frequency beating." The clearly audible phenomenon makes a listener aware of the non-symmetry of the field, potentially resulting in them taking a closer look at their data.

Discussion: Generally speaking, many idioms can be used for more than one search level, and which one they are used for seems to be dependent on the user's prior knowledge (both about the data and about the idiom). Nevertheless, the majority of papers in the surveyed literature offer the search level of "explore" and, in general, most papers use the same search level for both their visualization and their sonification parts (see Table 2 and Table 4). Some examples of designs employing the same search level are [CB17, DCM*18, ERI*22, FBC12, Rib19, GR11, HK22], while fewer combine different search levels for the two modalities [HAR16, RFK*15, RJ16, AK11, Bal15, CWM21].

Sonification	Visualization			
	ratio	interval	ordinal	nominal
ratio	21	11	0	10
interval	7	22	2	12
ordinal	0	2	8	7
nominal	4	2	5	12

Table 6: *The number of entries for the different levels of measurement and their combinations between the two display techniques. Most idioms use identical levels of measurement. Notably, 42 cases use a higher level of measurement with their sonification than with their visualization (numbers above the main diagonal). Twenty cases use the levels of measurement the other way around (numbers under the main diagonal).*
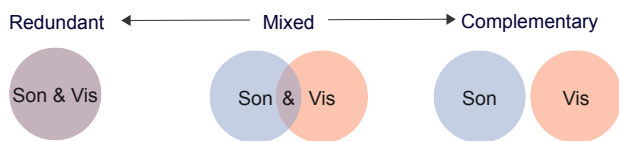
Figure 6: *The three levels of mapping-redundancy: Redundant mappings display the same information via visualization and sonification. Complementary mappings display different information to the two senses, and mixed mappings map some information redundantly and others in a complementary way.*

3.6. Dataset Type and Level of Measurement

Two more categories that we tagged individually for the sonification parts and the visualization parts are the dataset type (see Figure 5 and Table 2) and the level of measurement (see Table 2). We distinguish between the four dataset types of tables, networks, fields, and geometries [Mun15] and between the four levels of measurement of nominal, ordinal, interval, and ratio scale.

It is reasonable that most idioms display data from identical dataset types with both their visualization and their sonification. The most prominent dataset type is the table, used in designs such as [BTB23, DLVDCG22, DCM*18, ERI*22, FBC12, HAR16, MMM18, RJ16]. Table 5 reveals a notable exception from the dominance of identical dataset types: the combination of visualized geometry in combination with a sonified tables, often related to maps and supplemented regional information [BF12, LSB*23, MMU16, NRL*12, RS22].

Concerning the levels of measurement of the displayed data, we observe, again, that most idioms display data from the same level (see Table 5). We list a selection of cases for ratio data [ASH*12, DCM*18, GRK*16, PFH*22, Rön21], for interval data [BB19, DLVDCG22, Rib19, Her20, MMU16], for ordinal data [AK11, LF21, PCB23], and for nominal data [GKW21, HAR16, NSC16], and now will continue with a discussion of three methods to map the data to the senses, the levels of redundancy.

3.7. Level of Redundancy

When combining sonification and visualization, designers have three options – displayed in Figure 6 – to distribute the information they want to transport to the senses. They can have the same information represented in both the sonification and the visualization (redundant mappings), they can display one part of the data exclusively visually and another exclusively auditorily (complementary mappings), or they can map some parts of the data redundantly and some complementary (mixed mappings). While this continuum seems to be an intuitive description of audiovisual display idioms, we should distinguish between "technical redundancy" and "communicative redundancy."

Technical redundancy describes the technical mapping from the data to a visual and auditory representation. If all the displayed information is mapped to channels of both modalities, then the audiovisual display idiom employs a technically redundant mapping. An example is the auditory line graph: We see and hear identical information via spatial position (visualization) and pitch (sonification). In some cases, it is not enough to use a purely technical description of redundancy without incorporating our way of perceiving the different displays as humans. Therefore, we introduce the term "communicative redundancy."

Communicative redundancy describes the fact that technically redundant designs might encode the same information with different perceptual qualities. Hence, we could identify different patterns in data by listening to them than by looking at them. Communicative redundancy will usually be strongly related to technical redundancy, but there are exceptions, such as the combination of a WAV form visualization (line chart) and an audification, such as in the Sonification of the Tohoku Earthquake in Japan in 2011, also described in [WW15]. Technically, those two are fully redundant (the same data attributes are displayed both visually and auditorily), but communicatively, they complement each other.

In the following, when we speak of redundancy, we use the definition of technical redundancy. Communicative redundancy is most likely also dependent on the individual receiver, which is why we would not be able to consistently assign a tag to each case. Nevertheless, when implementing an audiovisual display idiom, a designer should consider its communicative redundancy just as well.

Rönnberg and Johannson [RJ16] present a technically redundant mapping of the density of a scatterplot or a parallel coordinates plot to auditory channels. Users would be able to hover over the plot and listen to an auditory representation of the density in the data. It is technically redundant as we are generally able to see the density of the plot by looking at it. In a communicative sense, the design is not redundant, which is why their evaluation shows that the additional sonification helps users identify especially dense areas. The eyes' ability to assess the visual density is supported by the sound of their design.

Dedicated environments that offer the design of audiovisual display idioms frequently employ redundant mappings [PC19, Rib18, CWM21, DLVDCG22, KLTW17, Rib19]. One of them is the *Sonigrapher* (see Figure 3(c)), presented by García [Rib19], that combines a line chart and a parameter mapping sonification of the light curves from NASA's exoplanet archive. The vertical position of a line is

mapped to the pitch of a musical sound, essentially "playing back" the line from left to right (compare Figure 7).

The work presented by Rau et al. [RFK*15] complementarily enhances the visualization of molecular simulations by using auditory icons and parameter mapping sonification. A user can position a "virtual microphone" inside a 3D-rendered visual representation of a molecule and listen to visually occluded processes.

Temor et al. [TMN*21] presented a mixed mapping approach for the audiovisual analysis of computational fluid dynamics tailored towards the prediction of cerebral aneurysm ruptures, hence in the medical context. Their sonification design is psychoacoustically motivated in a way that amplifies the differences between different simulations of their spatiotemporally dynamic data. The authors explicitly mention their choice to apply a mixed mapping in their design. They motivate their decision by the observation that the "*spatiotemporal fluctuations are highlighted in a way that seems to be superior to the presentation of different information to different sensory modalities, which is in line with how we make sense of spatiotemporally-dynamic stimuli in everyday scenarios.*" We consider this observation highly plausible and inspirational for potential future research and will reflect upon it in the concluding discussion of this STAR.

Maças & Martins, and Machado present an audiovisual display idiom that displays consumption patterns collected from Portuguese supermarkets over the course of two years. The authors, just in line with Temor et al. [TMN*21], also argue for the employment of a mixed design, their teams having made prior experience with complementary designs that seem to have been less effective.

Discussion: Regarding the level of mapping redundancy, a high-level observation is the following: When designing an audiovisual display idiom, mixed mapping seems to show a special potential. The mapping overlap between the sonification and the visualization seems to help the user perceive an idiom as integrated rather than as two displays existing next to each other. There are different options to design a mixed mapping, out of which a promising one seems to be to synchronize a visual animation with a sonification [TMN*21, MNW*18]. In a similar manner, synchronizing the spatial position of the visual display with the direction that a user perceives the sound from can be helpful to perceptually integrate the two stimuli [CB17]. An audiovisual analysis idiom that utilizes complementary or mixed mappings enables a designer to choose the employment of sound instead of a second view. Data that would conventionally be made visible via a second view can be represented using sound instead, which can modify the design of the visual view itself. Such situations can become especially apparent whenever the visualization needs to fit on small screens such as on smartwatches.

3.8. Evaluation Approaches

The evaluation of designs is a pressing issue in both the sonification and visualization fields. Audiovisual idioms likewise need to be evaluated. Rönnberg and Forsell even argue for the standardization of questionnaires that assess the usability of audiovisual representations that could be used in combination with other measurements [RF22]. To study the current practice of evaluation of audiovisual designs, we applied four of the classes of evaluation

techniques suggested by Isenberg et al. [IIC*13]. The four classes are user performance (UP), user experience (UE), algorithmic performance (AP), and qualitative result inspection (QRI).

When evaluations measure how specific features in a visualization or a sonification affect user performance with a system, these approaches belong to the evaluation class UP [IIC*13]. Controlled experiments using various time measurements and accuracy are typical example methods in this class. Rönnberg and Johansson [RJ16] explored the combination of visualization and sonification, where the users had to identify the visual area with the highest density in visual representations with and without the support of sonification. The authors measured accuracy and response time, and the results showed that the combination of visualization and sonification increased accuracy in comparison with visualization only but also that response times were longer when sonification was used.

The class of UE includes evaluations that focus on subjective feedback on and experiences with a visualization or a sonification [IIC*13]. Interviews and/or various questionnaires are common evaluation methods. In our corpus of papers, this was a common approach for evaluating the audiovisual representations and was often used in combination with other evaluation approaches. The paper by Ballweg et al., [BBV16] presents a study where users answered a questionnaire about their experience of an audiovisual system for drug design on a 5-point Likert scale. By comparing responses given before and after the study, it could be determined how the system could be integrated into the users' workflow, if sonification could have a positive effect on solving the task, and in what way the system could be improved.

Few papers in the systematic literature review used AP as an evaluation approach. Evaluation approaches in this class should contain a quantitative study of the performance or quality of visualization and/or sonification algorithms [IIC*13]. However, in our corpus of papers, this approach was only found in a few cases and was employed to determine that an audiovisual algorithm could produce sufficient quality rather than determining a certain level of quality or comparing different algorithms. As an example, in the paper by Kariyado et al. [KAV21], it is stated that the evaluation showed that the system presented in the paper allowed for a default amount of 255 audio sources with any audio drop-outs.

Evaluations in the class QRI aim to draw conclusions based on qualitative discussions and assessments of audiovisual representations [IIC*13]. In contrast to UE, these types of evaluations do not involve end users but instead, ask the user to assess the representation for themselves. In the work of Bru et al. [BTB23], an approach to a combined audiovisual representation based on parallel coordinates or dense line charts is presented. Several different attributes and characteristics in the representations are discussed based on the researchers' own reflections, but no formal user evaluation has been performed.

Discussion: We have now demonstrated there are various approaches for evaluating audiovisual representations. The aim of the work and the research questions that are asked determine the class of evaluation chosen. In some cases, in the corpus of papers, two of these evaluation approaches have been combined to provide a better and more detailed analysis of the outcome of the study

findings [BBV16, GRK*16, PFH*22]. Yet about a third of the studies included in the systematic review did not include an evaluation at all. The absence of evaluation might lead to a situation where promising audiovisual design ideas might be rejected because clear and convincing evidence of their usability, benefits, and function is not presented in a paper, or less useful ideas are overrated and promoted.

3.9. Target Platforms and Interactivity

Audiovisual display idioms can be displayed on different target platforms and in different environments (see examples in Figure 7). A typical display for the combination of sonification and visualization is the computer desktop environment, as this is also the most commonly used environment for visualization-only designs. About two-thirds of the designs included in this survey are developed for a desktop environment.

The second largest category of target platforms was physical and/or multi-user environments. This includes dome theaters and planetariums, which are commonly used for audiovisual display idioms toward topics of astronomy (as previously mentioned in subsection 3.1) [HTHB22, RS22, EEBR21]. Another type of physical environment is a dedicated immersive environment for collaborative data analysis, where several users can take part in the data exploration. One example of such an environment is the *CRAIVE-Lab* (Collaborative-Research Augmented Immersive Virtual Environment Laboratory), where a design visualized and sonified market data of 128 corporations by using the large panoramic display and the 128-channel loudspeaker array of the environment [CB17] (see Figure 7(b)).

A low-cost alternative for physical environments is a virtual one, which is most commonly enabled through head-mounted displays. An approach for an audiovisual display idiom in this type of environment is to let the user navigate the dataset from a first-person perspective and use spatialized sonification, which dynamically changes depending on where the user is positioned in the dataset. An example is the design of Berger & Bill (see Figure 7(a)), which facilitates an immersive exploration of urban noise standards [BB19] by creating a virtual environment of a city where the sonification allows the user to listen to the collected noise levels by navigating the environment.

The use of touch displays and other tangible interfaces for an audiovisual display idiom can allow for a unified integration of sonification and visualization. A recurring approach is to use the visualization as a graphical user interface that enables the sonification upon interaction. Through a touch display, the design by Ferguson et al. [FBC12] enables the user to filter a line graph with multi-touch gestures to select what data should be sonified. As a more analog approach, the system *Sonophenology* by Ness et al. [NRL*12] lets the user select areas on a printed color-coded paper map through fiducial markers, which in turn selects what information should be conveyed by the sonification.

About two-thirds of the papers in the database include some form of interaction. One of the most recurring objectives of the interaction was to make a selection of the dataset that would be sonified [FBC12]. This can include selecting a geographic region

to which the dataset is attributed to [NRL*12]. For example, the audiovisual display idiom of Matsubara et al. [MMU16] creates an interactive sonification system for exploration of seismic data through a horizontal and vertical range slider to specify a geographic region. Other forms of interaction include simpler tasks such as browsing sonifications of individual or groups of data objects [Rib19, LSB*23]. Huppenkothen et al. [HPDW23] created an interactive multimedia version of the Hertzsprung-Russell Diagram, where a user can select individual data entries of the diagram to listen to their sonification. On the other hand, there also exist more complex forms of interactions such as navigating a 3D environment [BB19, JMP13, MNW*18].

Papachristodoulou et al. [PBM15] created a design for navigating complex datasets of brain networks, where the sonification conveyed information about different brain regions of the dataset. Within the sonification community, it was the introduction of model-based sonification [Her02], that made user interaction an integral part of the data analysis process. An example of a model-based sonification is the *Mode Explorer* by Yang and Hermann [YH18], where scratching-interactions with a pencil on a scatter plot enable the user to investigate different modes and their properties through the sonification.

Discussion: When displaying an audiovisual display idiom on a target platform other than a desktop environment, it often creates opportunities for interacting with the system in a novel manner [EEBR21, HK22, NRL*12, PBM15]. In the context of data analysis, interaction is also a particularly relevant part of the user experience [BBV16, Bea11, EEBR21, HK22].

3.10. Users and Goals

Just as with visualization-only designs, audiovisual display idioms are often developed toward specific users and end goals. The most commonly occurring goal for the papers in the database was "data analysis." This could either be to present a dataset to the user or letting the user explore one or several datasets with the idiom, as described in subsection 3.2. The benefit of using visualization and sonification in this regard is that the user can gain different perspectives of the data through the two sensory modalities. For example, Alonso-Arevalo et al. [ASH*12] created a multimodal interface for curve shapes and curvature, where the user can evaluate the quality of a three-dimensional shape by using both the visual and auditory, and in this case even their haptic perception. The curve shape was mapped to the fundamental frequency of different carrier sounds, and the idiom offered different sound designs to convey the information.

The most recurring pair of users and goals in the database was data analysis for domain experts, which most commonly would involve an audiovisual display idiom to convey data in a specific application domain in the natural sciences. Temor et al. [TMN*21] and MacDonald et al. [MNW*18] created an auditory complement to a scientific visualization of computational fluid dynamics, where the sonification aided in understanding the temporal changes in the visual animation. Gionfrida et al. [GRK*16] suggest combining the visualization of brain scans (in the context of Alzheimer's dementia research) with a parameter mapping sonification that they call *Triple-Tone Sonification*. The design makes use of the fact that the human
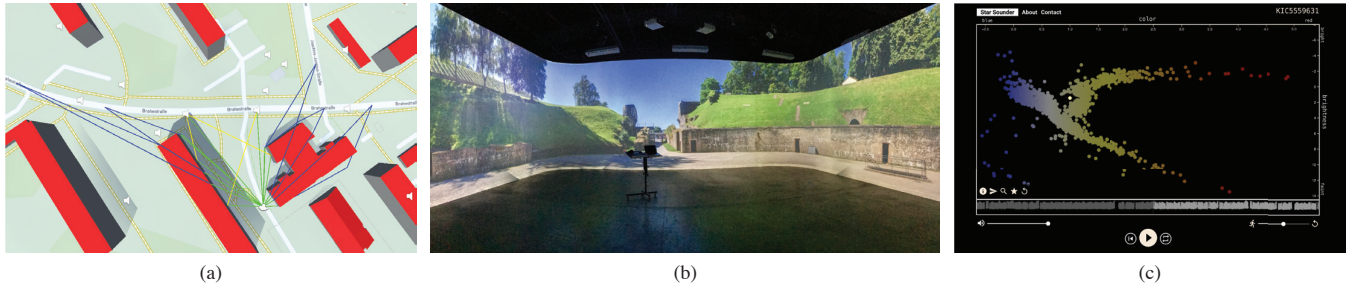
Figure 7: *Examples of different target platforms and environments. (a) The combination of VR visualization and sonification for an immersive exploration of noise in urban environments by Berger and Bill [BB19]. (b) An immersive virtual environment for audiovisual spatialized data sonification presented by Chabot and Braasch [CB17]. (c) A screenshot showing the interactive interface for the sonified Hertzsprung-Russel diagram by Huppenkothen et al. [HPDW23]. All screenshots CC BY 4.0.*

ear will perceive the sound of two or more very similar frequencies as "frequency beating," which is also employed in [MAFP19a]. For the purpose of process execution data analysis, Hildebrandt et al. [HAR16] demonstrated how incorporating sound could enhance the process of identifying anomalies or conducting root cause analysis for irregularities and errors. A design for a more general user case was created by North et al. [NSC16], which created a an idiom to convey Git version control data. The temporal nature of the data lent itself well to be sonified by sequentially going through the data, where an earcon is played when a commit occurs, and drum sounds are played whenever a conflict occurs in the data.

The second most commonly occurring pair of users and goals was public engagement for the general public, where the use of an audiovisual display idiom has the potential to enable higher engagement by conveying information through both the visual and auditory modality. This would commonly be targeted towards popular science topics such as climate change [HK22, Bal15] and astronomy [HTHB22, TB23]. As previously mentioned in subsection 3.2, the work by Russo et al. [RS22] is an example of a sonification targeted towards the general public for public engagement, where the response of the resulting videos on social media was in part used as a metric for the public engagement of the design.

Discussion: Regarding the successful design of audiovisual idioms that support their users in analyzing data, it is most relevant to include them in the design process. Both the visualization community and the sonification community have individually studied the relevance of including domain experts in their design process [SMM12, Gou17], and the same will be necessary for the integration of sonification and visualization. The co-author network shown in Figure 9 displays quite many domain experts being collaborators in the surveyed literature, which can be regarded as a promising sign in general.

4. Survey Data Analyses

So far, we have focused on the discussion of the audiovisual idiom designs themselves. In this section, we want to take an even broader perspective on the collected data. We will study existing relationships between different tags along a correlation matrix in subsection 4.1, as well as the network of co-authors and its implication for the field in subsection 4.2.

4.1. Correlation Matrix Analysis

To understand potential relationships between different tags within our classification, we computed a correlation matrix, shown in Figure 8. We want to highlight some of the found correlations in the data out of which most are not surprising, nevertheless, they are not obvious and our data enables us to take such a meta-perspective.

Close to the main diagonal of the correlation matrix, we highlighted four smaller matrices (**1**), showing an interesting phenomenon: Whenever the tagging of the papers was done individually for the visualization and the sonification parts, the two were mostly tagged with the same label. Hence, the visualizations and the sonification, in many cases, use the same *reading levels* [Ber83], the same *search levels* [Mun15], the same *dataset types*, and the same *levels of measurement*. This finding can motivate future research regarding the reading levels and the search levels. Studies could investigate the potential of other distributions of the reading level onto the senses.

Marked with (**2**), we highlighted four phenomena concerning the purpose of an audiovisual idiom. Idioms designed to present rather than explore data are likely to also use the reading level "whole" for their sonification ($r = 0.41$). On the other hand, idioms that are designed for exploration likely use the group reading level ($r = 0.3$). We see a correlation between the purpose of an idiom and its interactivity, with idioms for exploration more likely to be interactive ($r = 0.23$) and idioms for presentation more likely not to be interactive ($r = -0.48$). Also, domain experts use more idioms for exploration ($r = 0.44$), while the general public is more likely to be 'just' presented with data ($r = 0.50$), going hand in hand with their general goals of data analysis ($r = 0.52$) or public engagement ($r = 0.50$).

Two more relationships with the interaction tag emerge from the correlation matrix (**3**). A sonification that uses the reading levels of "group" or "single" is likely to offer user interaction ($r_{group} = 0.45$ and $r_{single} = 0.24$), while a sonification with the reading level of "whole" does not usually offer user interaction ($r_{whole} = -0.28$). Furthermore, designs that map data to the senses in a complementary way are more likely to offer interaction to their users ($r = 0.32$).

Figure 8: *The correlation matrix of a selection of classification tags. A diverging color palette is used with white in the center representing no correlation, teal (up) increasingly positive correlations, and brown (down) increasingly negative correlations. (1) reveals similar tags for visualization and sonification parts; (2) shows correlations between purpose, reading level, interactivity, and intended audience; (3) uncovers relationships with the interaction tag. A high-resolution version of this figure is part of the supplemental material.*

Overall, most of the computed correlation values in the correlation matrix are close to zero. Also, some of the existing higher values are likely not significant due to the low amount of data for the respective tags (such as with the AP evaluation, which has been used in only two papers). Therefore, we only discussed three perspectives that have considerable correlation values, are plausible, and that we consider most relevant to the bigger picture.

4.2. Co-Author Network and Development Over Time

As part of our meta-analyses, we studied the co-author network within our database, providing a good picture of the state of the research community interested in audiovisual analytics. Each node in the co-author network represents an author with a total of 165 nodes (in 48 teams), and a link connects two nodes if the corresponding authors have collaboratively contributed to a publication. We assigned a color to each node using the primary discipline of the author and classified them, to the best of our abilities, into one of four disciplines: sonification, visualization, domain, and unclassified. For each author, we considered their main publication focus, particularly around the time of their contribution to the works in our database, along with their background, education, their role and input to the relevant papers. Authors classified under "domain" are recognized as domain experts, and the category "unclassified" encompasses scenarios where the author's primary discipline does not fall into

Figure 9: *The co-author network, displaying 165 authors of the 57 papers in the database. The links represent papers written by a team of co-authors, and the colors display their primary disciplines. An interactive version of the network can be explored using the Orange workflow, which is part of the supplemental material.*

any of the other three categories, the author is in the early stages of their career without a defined research focus, or that the information about the author is not available.

The network offers several insights into the structure of the field. The majority of co-author teams are individual teams without interconnections. Only in four cases have authors collaborated with different (groups of) co-authors, usually because one person worked with different collaborators. This phenomenon differs from networks in other fields, such as the one described in a recent STAR on the use of embeddings in visual analytics [HWKK23], showing a network with clusters up to the size of 75 authors (note that the STAR on embeddings has 122 entries). We interpret the dominance of many disjoint groups as a sign for the rather young field and hope, with this STAR, to contribute to the future development of the field growing closer together. We identified 14 teams or individual authors contributing to the corpus that consist of sonification experts exclusively. Two of the teams consist only of visualization researchers, and 21 of the teams are collaborations between the communities, including domain experts in many cases. The inclusion of domain experts and the quite large number of diverse groups also including domain experts can be considered a promising sign for the future development of the field.

Our data also enables us to take a basic temporal perspective on the development of the field. During the years 2011 and 2023, overall, 57 papers were published. The majority of papers were published in the second half of this time span, pointing towards a generally increasing trend. A drop in publications in the year 2020 is likely a phenomenon related to the COVID-19 pandemic, being dominant worldwide, especially during the year 2020.

5. Adjacent Topics

As mentioned earlier, three distinct topics are adjacent to the scope of our STAR: audiovisual idioms in the context of accessibility, real-time monitoring, and arts. Although the three areas are related, they might be better served by a classification system different from the one we used. Still, they are relevant and can be inspirational for our field. The following three sections present work from those adjacent fields that have been curated by our team of co-authors. The sections do not claim completeness or the usage of a systematic approach but are intended as a general introduction for the interested reader.

5.1. Accessibility

Even though visualization is one of the most common ways of communicating data, many visualizations are inaccessible to readers with visual impairments. Some studies have suggested the use of sonification to support readers with visual impairments and a recent STAR on accessibility research in visualization identified 16 out of 56 papers that utilize sonification [KJRK21]. During the filtering and classification of our assembled literature, a category emerged exploring sonification for accessibility of visualization. The literature that was sorted into this category focused on the design and evaluation of sonification to improve the accessibility of visualizations for visually impaired people. As the literature in this category did not explore integrated audiovisual representations for data analysis in

general but rather used sonification as a support technique, these are not included in the systematic part of our STAR.

Sonification for accessibility in this way often suggests sonification designs leaning towards auditory graphs. Auditory graphs are the auditory equivalent to visual representations, such as plots, graphs, and charts [HHN11], but instead of mapping data properties to line positions or the size of an area, they are mapped to various auditory parameters such as pitch or loudness. Auditory graphs are suggested to be useful as sensory substitution and as a means of presenting information when line of sight is not possible [SNH05], and using sound to represent data makes data analysis more accessible [WM10]. To give the reader a brief introduction to this category, here are a few recent studies that have explored sonification for accessibility in visualization.

One study explored the accessibility of data visualizations presenting data to blind and visually impaired users [FFSR*23]. It was found that sonification might make discerning trends in the data more accessible, but also that the lack of experience in using sonification could lead to misinterpretations of the presented data. The study was conducted using screen readers, and the users often used alt-text information to support and validate what was heard through the sonification. Another study explored the use of natural sounds mapped to data visualizations, bar charts, and line charts [HEUHEB23]. The reason for using natural sounds was to support users without musical training in understanding the sonification and auditory graph, and it was found that these natural sounds could support the understanding of categorical data and were most useful for users without musical training. Infographics, the combination of visualization and text information, has been explored with an auditory-only approach [HGI*22]. An interactive approach, infosonic, was explored to facilitate accessibility of data to blind and low-vision users, using spoken introduction and annotation, and non-speech sonification. The study shows that the sonification approach supported understanding and forming a mental image of the data.

Often, accessibility studies tend to blend sonification, i.e., non-speech sounds, with generated speech or screen readers. This makes the interpretation of the results from these studies somewhat challenging. The findings, i.e., the experiences and the understanding of the sonification, might be due to the sonification itself or depending on other auditory stimuli used. Also, the visual representation, the user interaction, and the user interface are of importance to the study outcome. In visual data analysis, interaction is essential for exploring the data [CGM19], and similarly, for sonification to be a useful tool, dynamic human interaction is necessary [HH04]. Sonification for accessibility of visualization for visually impaired users is an exciting approach and some of these findings could suggest interesting approaches also for sighted users in supporting visual perception with sonification [Rön19a, Rön19b], or reducing cognitive load [ZPR16] on the visual modality by sonification (see for example [MLS95]).

5.2. Monitoring

A comprehensive overview or a systematic literature review of audiovisual process monitoring is far beyond the scope of this report.

This section is intended to provide some insights into this research field and to distinguish the various scientific communities involved.

The combination of visual and auditory displays for process monitoring purposes has been well-established in real-world applications and interfaces. The audiovisual representation of Morse code for an exchange of information serves as an early example of a multimodal monitoring display and interface. In supervising and controlling the states of various devices and processes, audiovisual interfaces combine the advantages of both modes: volatile alerting sound signals enhance situation awareness, while visual cues provide additional information about the situation, enabling users to take action. Implementations of audiovisual interfaces are ubiquitous in both professional and day-life environments. For instance, consider the warning beep, which can be heard when the temperature outside your car drops below 4°C, alerting you to the possibility of slippery road conditions. A quick glance at the visual display will provide you with the exact temperature and the opportunity to decide whether to take action or not.

Besides the approaches to audiovisual data analysis described in the previous sections of this report, audiovisual monitoring interfaces play an integral role in various domains, including air traffic control [EBL*23], control rooms [SFLD22, HHG*12], anesthesia [AQH*21, RMS*22], neurology [LCS*28], dermoscopy [WRK*19], surgery [Zie23], network monitoring [Wor19a, AHvR*21], automotive [JDS15, XWX*22], to name a few.

During the literature search for this STAR, several publications were labeled with the tag "monitoring." After closer investigation, which involved excluding multiple publications written by identical authors on very similar topics, as well as publications addressing topics of accessibility and artistic media installations, only a relatively small representative subset of 14 findings qualifies for audiovisual process monitoring within the search criteria of this STAR.

The reasons for this rather limited number of retrieved articles in this category are manifold. Firstly, most research on monitoring applications used for auditory feedback of human behavior in the fields of sports, therapy, or rehabilitation focuses on the design, impact, and evaluation of interfaces that utilize the non-intrusive attributes of auditory displays. Hence, they enable users to perceive information without any interfering actions, such as adjusting their head or body position to view a visual display [SJMT19, VRGM20, MNS16]. While some of these interfaces include visualizations at a low scale level, they hardly qualify as audiovisual monitoring devices in terms of providing a balanced contribution from both modalities.

In industrial and surveillance contexts, factors such as cognitive workload, perception organization, situational awareness, alarm fatigue, and deafness play crucial roles in the successful implementation of infrastructures. However, these attributes are discussed in other scientific communities, including Human Factors and Applied Ergonomics, as well as several areas of medicine and health, particularly in anesthesia, rehabilitation, sports, and cognitive psychology. In addition to performance comparisons between auditory, visual, and audiovisual modalities [AG19], particular interest is given to the dual-task paradigm, which evaluates, for instance, the ability to identify an event requiring action in a secondary domain while simultaneously performing a primary task [LWP*13, HHRM16, TML*15, NGJW14].

(Non-)systematic literature reviews have been published in several of the mentioned areas [WMY*17, KIK19, SMS21, HRM15]. To the best knowledge of the authors, no comprehensive STAR covering the entire spectrum of audiovisual or related multimodal monitoring applications and approaches has been conducted thus far. However, there are some fundamental publications treating the specific attributes and criteria for research and development in the field, such as [Ibe20, Joh04, WS07].

5.3. Arts

During the classification of the corpus, an "art" category emerged, defining projects combining visualization and sonification for artistic purposes. We understand the emergence of art practices that intertwine sonification and visualization not only as inspiring for designers and audiences alike but also relevant to the broader discourse around data representation towards a better, more efficient, and engaging human-data relationship.

Working with the literature, we understood artistic endeavors call for a different kind of classification than idioms intended for audiovisual data analysis. We decided not to include such contributions in the systematic part of this STAR and excluded papers where authors made their intentions of an artistic contribution explicit in the paper.

Nevertheless, to give the reader a brief introduction to the field, we present representative cases of an emerging art category that could form the basis of future research on the integration of sonification and visualization. In addition to the artistic cases that emerged from our literature search, we use four more sources to curate a collection of representative cases for artistic endeavors: The Data Sonification Archive (DSA), the Computer Music Journal, the Leonardo Music Journal, and the Ars Electronica Festival. The Data Sonification Archive is an online crowd-sourced collection of data sonification projects. Out of 455 cases currently hosted in the DSA, two of the co-authors of this report recently categorized 139 cases as "art", i.e., as having their primary purpose in creating and delivering an artistic experience to an audience [LLC23]. A preliminary search on some of the major venues for academic artistic publications – Computer Music Journal and Leonardo Music Journal – accounts for 182 and 118 cases, respectively, in line with the volume of art-related cases in the Data Sonification Archive. A search of the archive of the renowned Ars Electronica Festival on projects that translate data into an audiovisual idiom returns 32 cases, again, in line with the 23 projects currently hosted on the DSA.

Typically, data-driven artworks that use an audiovisual design address a non-expert audience and can take different forms, such as live performances, multimedia installations, video productions, or web experiences where the sonification and visualization of the same dataset are combined. The DSA classification highlights recurring topics such as climate change and/or the representation of environmental data, as well as individual and collective interaction in the urban space, internet, and social media. The predominance of socially relevant topics shows that the emerging art category presents borderline characteristics with the adjacent category (in the DSA classification system, see [LLC23]) of "public engagement".

This latter category identifies projects that, although being often presented in the form of an artistic experience, have explicitly and primarily the goal of increasing awareness – even fostering activism – among the general public on a specific phenomenon.

This is the case for projects that present data related to climate change, such as *Klima|Anlage—Performing climate data* [GHJ*19], one of the cases identified as artistic in our literature search. In this work, climate data and model predictions between 1950 and 2100 can be chosen by the listener for twelve selected regions of the world and interactively sonified and visualized in an immersive installation environment with the goal of contributing to "the urgent need to inform the general public about climate change".

In *Too Blue* [Foo15], hosted on the DSA, the author created an audiovisual experience (in the form of a video) where music generated by tracking the land loss in coastal Louisiana over 78 years due to man-made levees, drilling and dredging for oil and gas, and climate change is combined with an aerial color-coded visual map that shows the increase in the "blues" colored areas, corresponding to the sea, in the coastal land of Louisiana. In *Heat and the Heartbeat of the City* [Pol06], the urgency to communicate climate change to the broader audience combines with a reflection on the urban space through the simultaneous sonification and visualization of actual and projected data on increased temperature in New York City. In *aqua_forensic* [SŠ18] data on invisible pollutants in the Danube river's water, and the effect on its ecosystem, are collected, sonified, and visualized 'to present the results of this scientific research to the wide audience in a poetic and artistic way.' Other prominent social issues such as personal data collection and privacy in the age of social media are at the center of multi-awarded art projects such as *Digital Violence* [FE21] where sonification is explicitly used to increase the audience's engagement with the visualized dataset on the issue of illegal government surveillance practices. From the collective to the individual experience, projects such as *Orbuculum* [Ylm23], and *Deep Sync* [KOR23] invite the public to interactively engage with personal psychological and physiological dimensions to create an immersive auditory and visual experience. In a more-than-human perspective, works such as *Spider Web Sonification* [SQS*20] and *Biota Beats* [KGS*20] shift the focus from humans to other species in an effort to support the public engagement with science.

Data art can stimulate an "artistic affectivization" that can contribute to building shared perspectives within a community [BMLA22]. The engagement of the audience with the phenomenon behind the data, however, also implies an information and knowledge transfer [MVC*10], i.e., the public not only has to connect emotionally with the data but also has to understand their meaning. The emergence of a category of artistic projects that combine sonification and visualization with the explicit intent to activate the audience on a socially relevant issue shows that multimodal approaches have the potential to generate impact, not only for education and awareness but also for stimulating action.

6. Concluding Discussion and Future Work

In this STAR, we have provided an overview of the field of audiovisual idiom design tailored toward data exploration and presentation. We have used a variety of perspectives to classify the existing literature, and, in section 5, we offered insights into three topics that

are adjacent to our report and can be highly inspirational for future developments. For our categorizations, we also provided brief discussions or reflections as part of the individual subsections 3.2 to 3.10.

Before we adopt a global viewpoint and discuss broader insights, challenges, and exciting future research opportunities, we now want to mention some of the key insights and results from previous sections. Overall, our corpus holds a quite diverse field of combined designs, employing visual idioms such as scatter plots, maps, and networks, as well as sonification techniques such as parameter mapping sonification, audification, and earcons.

The analysis of employed reading levels [Ber83] of the different designs revealed an imbalance between the two modalities. The visualization part of the identified audiovisual idioms mostly works at higher reading levels than their sonification counterparts. This phenomenon suggests that many designs provide an overview of the data using their visualization, while the sonification supports the analysis of details.

Such design decisions are also related to the level of redundancy an audiovisual idiom employs, and mixed approaches, which represent parts of the data redundantly and other parts in a complementary manner, seem to show special potential. Such a selective alignment of mapping between the sonification and visualization might assist us in perceiving an idiom as integrated rather than as two separate displays coexisting next to each other.

Furthermore, the option to use sonification as part of an idiom's design can directly influence the visual view as well. Thinking of idioms with "multiple views," one of the views could, for example, be covered by an "auditory view." For visualization designers, such considerations are especially relevant when space is limited, as is the case, for example, on smartwatches [LDIC21].

The culture of providing supplemental material: A thriving future for the field of audiovisual idiom design also depends on the communities' culture when providing demos for their designs as supplemental material. Out of 57 articles in our surveyed literature, only 29 provided a demo that is still available today. Five provided a demo that is no longer available, and 23 provided no demo at all. Online repositories such as Zenodo, hosted by CERN, the platform OSF, maintained by the Center for Open Science, or long-term archives such as Phaidra, hosted by the University of Vienna, allow the upload of videos and even the assignment of DOIs. While making sound available in a paper used to be an explicit challenge to the sonification field, it is not the case anymore. An additional visual representation of the sound can also be included in the paper, using a spectrogram of the sound itself or an iconic representation of it.

Evaluating audiovisual idioms: It is often challenging to conduct systematic, well-informed, and, above all, properly performed and reported evaluations of audiovisual idioms, especially when the goal is to compare different studies. As mentioned earlier, Rönnberg and Forsell suggested the standardization of questionnaires tailored to the field in order to increase the comparability between different studies [RF22]. Establishing this comparability is particularly difficult because the two senses are used simultaneously in an audiovisual display, and the evaluation methods must adapt to this characteristic of the design. Specifically, the existing methods fo-

cus on several aspects. One is user performance, which studies quantitative metrics such as task completion times and precision of responses. Another is user experience, which focuses more on qualitative metrics such as user engagement and memorability. From the data we collected in our STAR, we can draw high-level conclusions regarding the type of evaluations that have been used within the corpus of literature.

Among the papers that included a user evaluation (21 of the 57 papers), most reported a subjectively stated benefit of sonification in perceiving and interpreting the visualization. Some subjective ratings and feedback also suggested that the use of sonification was engaging (six papers considered this [HK22, RS22, EEBR21, KAV21, Rön21, DCM*18]), could provide emotional content (mainly two papers [RS22, Rön21]), was experienced as interesting (five papers [HK22, RS22, EEBR21, Rön21, DCM*18]), as well as pleasant (four papers [RS22, EEBR21, Rön21, DCM*18]), and in perceived improved quality of visual information (six papers [PFH*22, EEBR21, KAV21, FN18, RJ16, BF12]). Other papers also reported subjective and objective measurements of improved interpretation and higher understanding of data when sonification and visualization were used in combination (16 papers [PFH*22, CWM21, EEBR21, DCM*18, FN18, GDAS*18, YH18, BBV16, GRK*16, NSC16, RJ16, PBM15, PBV14, ASH*12, BF12, Bea11]). This was shown in terms of higher scores or increased accuracy when sonification was used. However, this increase in accuracy was reported to be related to increased task completion times (see, for example, in [RJ16]). Similar effects have also been shown in other studies not included in this systematic review [MBKSM16, Rön19a, Rön19b]. The case might be that the use of the additional information provided by sonification that makes the increased accuracy possible takes a longer time to process and assess. While traditionally increased task completion times are considered a disadvantage, in some contexts, this can be seen as positive. Slowing down the user, resulting in them spending more time with their data, studying it from more and from different perspectives (including the auditory perspective), might reveal structures or patterns in a dataset that would have stayed silent and unseen if analyzed with a conventional unimodal display [HAS11].

These results indicate that there is, in general, a benefit of combining visualization and sonification for data exploration and analysis, regardless of the specific visualization and/or sonification design. Nevertheless, we consider it a major task for the future research of our community to systematically study what kinds of benefits the integration of sonification and visualization can provide.

The potential of unconventional task distribution: The data from our STAR shows that reading levels [Ber83] are not equally distributed to the two senses: The vast majority – 50 out of 57 papers used the reading level "whole" for their visualization (see Table 2), while only about half of the papers use the same reading level for their sonification parts. This observation might be directly related to the way audiovisual idioms distribute the tasks of overview and detail to the two senses. The currently more popular distribution seems to be to use visualization to display an overview of the data, while sonification is employed to provide details. While this distribution seems like an intuitive choice, future research should study the potential of switching the roles between sonification and visualization.

It is our daily lives and the way we experience the world around us that suggests sonifying an overview and visualizing details. Our auditory sense constantly screens our 360° environment while our visual sense actively focuses on details in our environment. The field of auditory process monitoring makes use of exactly this situation when alerting us of a status, regardless of our current visual focus. Nevertheless, we identified only three cases where the reading level for the sonification was "whole" and the one for the visualization was "group" [DLVDCG22, FBC12, YH18].

A second phenomenon we observe in the data is that many audiovisual display idioms employ the same search level [Mun15] (see Figure 8), and most of them employ the search level of "explore" (see Table 2). Again, it seems like a fruitful future endeavor to study the opportunities of designs that explicitly break such a pattern.

The potential of integrating sonification and scientific visualization: Scientific visualization typically displays inherently spatial phenomena and is often used to represent data that varies over time. The inherently spatio-temporal quality of such displays constitutes their potential to be integrated with sonification, as sonification is a display technique tailored toward temporal data structures [GGB05, ERI*23]. We relate this argument back to the real-world example from our introduction, where the combination of our visual and auditory senses offers a better understanding of a phenomenon than each sense alone: Rain falling outside of a closed window. Opening the window to not only see but also hear the sound of the rain gives us a better estimation of the amount and intensity of the rainfall. Abstractly, in such a situation, we perceive a spatio-temporal phenomenon. It is most plausible that adding sound to a time series of scientific visualization can increase its holistic interpretability in an ecologically valid manner [Neu04]. This phenomenon is shown in two publications in the database, in the work by MacDonald et al. [MNW*18], and the work by Temor et al. [TMN*21]. Similarly, these ideas also apply to animations in information visualization, where temporal aspects of the data might be visualized and sonified.

Design frameworks: When it comes to the possibilities of designing audiovisual display idioms, most researchers, as well as most domain experts, will not be able to develop their own designs. Interdisciplinary knowledge bridging visualization, sonification, interactive design, and human perception is necessary to design effective, engaging, and re-usable audiovisual display idioms. Therefore, we cannot expect a domain expert to be able to quickly draft a prototype the same way they may be able to do using established visualization-only frameworks such as *matplotlib* [Hun07].

In our database, we identified several design frameworks tackling this challenge [PC19, PCB23, LF21, DLVDCG22, KLTW17, CWM21]. Their primary focus is the design of the sonification part of the idiom, which is why we cannot speak of truly balanced contributions to both senses. A promising endeavor to enable truly balanced designs, with both the visualization and the sonification being equally well-informed, might be to not design them using one but two separate frameworks, each specialized for its purpose. In this regard, the recent work by Reinsch and Hermann [HR21, RH22, RH23] is promising, as it provides a sonification design framework embedded in the *Python* environment and is conceptually inspired by visualization design tools such as *matplotlib*. With both the sonifi-

cation design and the visualization design happening in the same environment, in this case, for example, a *Jupyter Notebook*, both can be developed with the required quality to meet their individual standards. Regarding the introduction of users to a new design environment, be it for visualization, sonification, or their combination, effective "onboarding" is a pressing issue. In the visualization community, the topic has been studied in recent years [SPA23, SWG*23], and the same will be necessary for sonification and audiovisual display idioms. Again, when it comes to sonification as a technique to display data, the prior experience of domain experts or the general public is low. This lack of experience can, in the short term, only be met with carefully designed onboarding processes.

Adjacent topics as inspiration to the field: Our brief look into the adjacent fields of monitoring, accessibility, and arts shows a special potential for inspiration for our communities' future work. The field of monitoring offers established evaluation methods that could potentially be adopted to support the evaluation of audiovisual idioms tailored toward exploratory data analysis. During our unstructured investigation into the field, we also identified the potential of a systematic STAR as a future contribution to the field of monitoring.

Also, the adjacent field of accessibility can inspire future research, where understanding the analysis patterns of visually impaired individuals can inform design decisions for idioms tailored towards sighted people [WM10]. While the field of accessibility can be informative to the field of audiovisual idioms, the same holds true the other way around. Especially the designs from our STAR that use redundant mapping to both the visual and the auditory senses can inspire future designs for visually impaired users. Furthermore, such designs have the potential to foster successful collaborative data analysis involving both sighted and blind users.

The identified artistic contributions show great potential to be inspirational to the field of audiovisual idiom design as well. In the future, effort should be made by the research community to develop a specific framework for the evaluation and analysis of audiovisual idioms at the crossing of art and public engagement. Such efforts need to explore how specific design strategies (e.g., interactivity and embodiment, as used in many of the cases presented in subsection 5.3) can be combined to support both affect and sense-making. Again, a systematic STAR dedicated to the artistic perspective would be a highly timely contribution to the community.

An audiovisual analytics community: With this state-of-the-art report, we hope to contribute to the future establishment of systematic research on audiovisual display idiom design. The current development concerning the studied co-author network and the temporal development of publication numbers in the field is promising. We hope to reach researchers from both the visualization and sonification communities, inspiring them for future interdisciplinary collaborations, and to *open their ears and take a look*.

Supplemental Material

As supplemental material to this STAR, we provide at https://phaidra. fhstp.ac.at/o:5541:

- corpus.bib and corpus.rdf: publication metadata of the surveyed literature in BibTeX and Zotero RDF format.

- corpus-tagging.csv: a table holding the surveyed literature and all of the used tags. Interested readers may use the table to identify articles that fall into specific combinations of classes.
- authors.csv: a table holding the names of the authors in our STAR corpus and their primary discipline. Interested readers may use it to identify potential future collaborators.
- corrmat.pdf: a high-resolution version of Figure 8.
- authornet.ows: the Orange Data Mining workflow file, used to generate the co-author network in Figure 9 (incl. the used color scheme file authornet_colorscheme.colors).

In addition, the corpus of relevant papers with tags, publication metadata, and full text of open access work is available as a *Zotero* library at https://www.zotero.org/groups/integrationsonificationvisualization/items.

References

[AG17] ANDREOPOULOU A., GOUDARZI V.: Reflections on the representation of women in the International Conferences on Auditory Displays (ICAD). In *Proceedings of the 23rd International Conference on Auditory Display (ICAD 2017)* (Pennsylvania State University, June 2017), Georgia Institute of Technology, pp. 43–48. doi:10.21785/icad2017.031. 4

[AG19] AUDRY E., GARCIA J.: Congruent audio-visual alarms for supervision tasks. In *Proceedings of the 25th International Conference on Auditory Display (ICAD 2019)* (Newcastle upon Tyne, June 2019), Department of Computer and Information Sciences, Northumbria University, pp. 7–11. doi:10.21785/icad2019.022. 20

[AG21] ANDREOPOULOU A., GOUDARZI V.: Sonification first: The role of icad in the advancement of sonification-related research. In *Proceedings of the 26th International Conference on Auditory Display* (June 2021), Georgia Institute of Technology, pp. 65–73. URL: http://hdl.handle.net/1853/66335. 4, 6

[AHvR*21] AXON L., HAPPA J., VAN RENSBURG A. J., GOLDSMITH M., CREESE S.: Sonification to support the monitoring tasks of security operations centres. *IEEE Transactions on Dependable and Secure Computing 18*, 3 (2021), 1227–1244. doi:10.1109/TDSC.2019. 2931557. 20

[AJB*18] ARBON R. E., JONES A. J., BRATHOLM L. A., MITCHELL T., GLOWACKI D. R.: Sonifying stochastic walks on biomolecular energy landscapes. In *Proc. 24th International Conference on Auditory Display (ICAD 2018)* (Michigan Technological University, June 2018), Georgia Institute of Technology, pp. 232–239. doi:10.21785/icad2018. 032. 7, 8

[AK11] ADHITYA S., KUUSKANKARE M.: The sonified urban masterplan (sum) tool: Sonification for urban planning and design. In *Proc. 17th International Conference on Auditory Display (ICAD-2011)* (Budapest, Hungary, June 2011). http://hdl.handle.net/1853/51918. 7, 8, 13, 14

[AQH*21] ANDRADE E., QUINLAN L., HARTE R., BYRNE D., FALLON E., KELLY M., CASEY S., KIRRANE F., O'CONNOR P., O'HORA D., SCULLY M., LAFFEY J., PLADYS P., BEUCHÉE A., ÓLAIGHIN G.: Augmenting critical care patient monitoring using wearable technology: Review of usability and human factors. *JMIR Human Factors 8*, 2 (2021), e16491. doi:10.2196/16491. 20

[ASH*12] ALONSO-AREVALO M. A., SHELLEY S., HERMES D., HOLLOWOOD J., PETTITT M., SHARPLES S., KOHLRAUSCH A.: Curve shape and curvature perception through interactive sonification. *ACM Trans. Appl. Percept. 9*, 4 (Oct. 2012). doi:10.1145/2355598.2355600. 7, 8, 14, 16, 22

[Bal15] BALLORA M.: Two examples of sonification for viewer engagement: Hurricanes and squirrel hibernation cycles. In *Proc. 21st International Conference on Auditory Display (ICAD-2015)* (Graz, Austria, July 2015), Georgia Institute of Technology, pp. 300–301. http://hdl.handle.net/1853/54172. 7, 8, 13, 17

[Bar97] BARRASS S.: *Auditory Information Design*. PhD thesis, Australian National University, Canberra, 1997. 3

[Bar12] BARRASS S.: The aesthetic turn in sonification towards a social and cultural medium. *AI & Society 27*, 2 (May 2012), 177–181. doi:10.1007/s00146-011-0335-5. 3

[BB12] BEARMAN N., BROWN E.: Who's sonifying data and how are they doing it? a comparison of ICAD and other venues since 2009. In *Proceedings of the 18th International Conference on Auditory Display* (2012), pp. 231–232. URL: http://hdl.handle.net/1853/44402. 6

[BB19] BERGER M., BILL R.: Combining vr visualization and sonification for immersive exploration of urban noise standards. *Multimodal Technologies and Interaction 3*, 2 (May 2019), 34:1–34:15. doi:10.3390/MTI3020034. 7, 8, 14, 16, 17

[BBV16] BALLWEG H., BRONOWSKA A. K., VICKERS P.: Interactive sonification for structural biology and structure-based drug design. In *Proceedings of ISon 2016, 5th Interactive Sonification Workshop, CITEC, Bielefeld University* (Bielefeld, Germany, 2016). 7, 8, 12, 15, 16, 22

[Bea11] BEARMAN N.: Using sound to represent uncertainty in future climate projections for the united kingdom. In *Proc. 17th International Conference on Auditory Display (ICAD-2011)* (Budapest, Hungary, June 2011). http://hdl.handle.net/1853/51922. 7, 8, 11, 16, 22

[Ber83] BERTIN J.: *Semiology of Graphics: Diagrams Networks Maps*. University of Wisconsin, Madison, 1983. Originally published in 1967 in French. 3, 6, 12, 17, 21, 22

[BF12] BEARMAN N., FISHER P. F.: Using sound to represent spatial data in arcgis. *Computers & Geosciences 46* (Sept. 2012), 157–163. doi:10.1016/j.cageo.2011.12.001. 7, 8, 11, 13, 14, 22

[Bly82] BLY S. A.: *Sound and computer information presentation*. PhD thesis, University of California, Davis, 1982. 3

[BM13] BREHMER M., MUNZNER T.: A multi-level typology of abstract visualization tasks. *IEEE Trans. Visualization and Computer Graphics 19*, 12 (2013), 2376–2385. doi:10.1109/TVCG.2013.124. 3

[BM20] BOUCHARA T., MONTÈS M.: Immersive sonification of protein surface. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (Mar. 2020), pp. 380–383. doi:10.1109/VRW50115.2020.00082. 7, 8

[BMLA22] BUENING R., MAEDA T., LIEW K., ARAMAKI E.: Between fact and fabrication: How visual art might nurture environmental consciousness. *Frontiers in Psychology 13* (2022). doi:10.3389/fpsyg.2022.925843. 21

[BPS06] BERTINI E., PLAISANT C., SANTUCCI G. (Eds.): *Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization*. ACM, 2006. doi:10.1145/1168149. 3

[BTB23] BRU E., TRAUTNER T., BRUCKNER S.: Line Harp: Importance-driven sonification for dense line charts. In *Proceedings of IEEE Visualization and Visual Analytics (VIS) 2023 – Short Papers* (2023), pp. 186–190. arXiv:2307.16589, doi:10.1109/VIS54172.2023.00046. 7, 8, 13, 14, 15

[CB17] CHABOT S., BRAASCH J.: An immersive virtual environment for congruent audio-visual spatialized data sonifications. In *Proc. 23rd International Conference on Auditory Display (ICAD–2017)* (Pennsylvania State University, June 2017), Georgia Institute of Technology, pp. 203–210. doi:10.21785/icad2017.072. 7, 8, 13, 15, 16, 17

[CGM19] CENEDA D., GSCHWANDTNER T., MIKSCH S.: A review of guidance approaches in visual data analysis: A multifocal perspective. *Computer Graphics Forum 38*, 3 (2019), 861–879. doi:10.1111/cgf.13730. 19

[CH04] COLLIER W. G., HUBBARD T. L.: Musical scales and brightness evaluations: Effects of pitch, direction, and scale mode. *Musicae Scientiae 8* (2004), 151–173. doi:10.1177/102986490400800203. 4

[CLR22] CAIOLA V., LENZI S., RICCÒ D.: Audiovisual sonifications: A design map for multisensory integration in data representation. In *DRS2022: Bilbao* (2022). doi:10.21606/drs.2022.380. 1, 4, 11, 12

[CMS99] CARD S. K., MACKINLAY J. D., SHNEIDERMAN B. (Eds.): *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann, San Francisco, 1999. 1, 3

[CWM21] CANTRELL S. J., WALKER B. N., MOSENG Ø.: Highcharts sonification studio: An online, open-source, extensible, and accessible data sonification tool. In *Proc. 26th International Conference on Auditory Display (ICAD 2021)* (Virtual Conference, June 2021), Georgia Institute of Technology, pp. 210–216. doi:10.21785/icad2021.005. 7, 8, 11, 13, 14, 22

[DB13] DUBUS G., BRESIN R.: A Systematic Review of Mapping Strategies for the Sonification of Physical Quantities. *PLoS ONE 8*, 12 (2013), e82491. doi:10.1371/journal.pone.0082491. 4, 11, 12

[dC07] DE CAMPO A.: Toward a data sonification design space map. In *Proceedings of the 13th International Conference on Auditory Display* (2007), pp. 342–347. URL: http://hdl.handle.net/1853/50042. 3

[DCM*18] DU M., CHOU J.-K., MA C., CHANDRASEGARAN S., MA K.-L.: Exploring the role of sound in augmenting visualization to enhance user engagement. In *2018 IEEE Pacific Visualization Symposium (PacificVis)* (Kobe, Apr. 2018), IEEE, pp. 225–229. doi:10.1109/PacificVis.2018.00036. 7, 8, 11, 13, 14, 22

[DLVDCG22] DE LA VEGA G., DOMINGUEZ L. M. E., CASADO J., GARCÍA B.: Sonouno web: An innovative user centred web interface. In *HCI International 2022 – Late Breaking Posters* (Cham, 2022), Stephanidis C., Antona M., Ntoa S., Salvendy G., (Eds.), Communications in Computer and Information Science, Springer Nature Switzerland, pp. 628–633. doi:10.1007/978-3-031-19679-9_79. 7, 8, 11, 14, 22

[EBL*23] ELMQUIST E., BOCK A., LUNDBERG J., YNNERMAN A., RÖNNBERG N.: SonAir: the design of a sonification of radar data for air traffic control. *Journal on Multimodal User Interfaces 17*, 3 (2023), 137–149. doi:10.1007/s12193-023-00404-x. 20

[EEBR21] ELMQUIST E., EJDBO M., BOCK A., RÖNNBERG N.: Openspace sonification: Complementing visualization of the solar system with sound. In *Proc. 26th International Conference on Auditory Display (ICAD 2021)* (Virtual Conference, June 2021), International Community for Auditory Display, pp. 135–142. doi:10.21785/icad2021.018. 7, 8, 11, 13, 16, 22

[ERI*22] ENGE K., RIND A., IBER M., HÖLDRICH R., AIGNER W.: Towards multimodal exploratory data analysis: Soniscope as a prototypical implementation. In *Proc. 24th Eurographics Conference on Visualization (EuroVis 2022) - Short Papers* (Rome, 2022), The Eurographics Association, pp. 67–71. doi:10.2312/evs.20221095. 7, 8, 11, 13, 14

[ERI*23] ENGE K., RIND A., IBER M., HÖLDRICH R., AIGNER W.: Towards a unified terminology for sonification and visualization. *Journal on Personal and Ubiquitous Computing 27* (2023), 1949–1963. doi:10.1007/s00779-023-01720-5. 1, 3, 11, 22

[FB18] FERGUSON J., BREWSTER S. A.: Investigating perceptual congruence between data and display dimensions in sonification. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (2018), ACM, p. 611:1–611:9. doi:10.1145/3173574.3174185. 4

[FBC12] FERGUSON S., BEILHARZ K., CALÒ C. A.: Navigation of interactive sonifications and visualisations of time-series data using multitouch computing. *Journal on Multimodal User Interfaces 5*, 3 (May

2012), 97–109. doi:10.1007/s12193-011-0075-3. 7, 8, 11, 13, 14, 16, 22

[FBT97] FLOWERS J. H., BUHMAN D. C., TURNAGE K. D.: Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples. *Human Factors 39* (1997), 341–351. doi:10.1518/001872097778827151. 4

[FBZ*99] FROHLICH B., BARRASS S., ZEHNER B., PLATE J., GOBEL M.: Exploring geo-scientific data in virtual environments. In *Proceedings Visualization '99 (Cat. No.99CB37067)* (1999), pp. 169–173. doi:10.1109/VISUAL.1999.809884. 4

[FE21] FORENSIC ARCHITECTURE, ENO B.: Digital violence, 2021. Webpage: https://sonification.design/#DigitalViolence. Last accessed: Dec 21st 2023. 21

[FFSR*23] FAN D., FAY SIU A., RAO H., KIM G. S.-H., VAZQUEZ X., GRECO L., O'MODHRAIN S., FOLLMER S.: The accessibility of data visualizations on the web for screen reader users: Practices and experiences during covid-19. *ACM Transactions on Accessible Computing 16*, 1 (2023), 1–29. 19

[FN18] FITZPATRICK J., NEFF F.: Stream segregation: Utilizing harmonic variance in auditory graphs. In *Proc. 15th Sound and Music Computing Conference (SMC2018)* (Limassol, Cyprus, 2018), Zenodo, pp. 52–59. doi:10.5281/zenodo.1422501. 7, 8, 11, 22

[Foo15] FOO B.: Too Blue, 2015. Webpage: https://sonification.design/#TooBlue. Last accessed: Dec 21st 2023. 21

[FPS*21] FRANCONERI S. L., PADILLA L. M., SHAH P., ZACKS J. M., HULLMAN J.: The science of visual data communication: What works. *Psychological Science in the public interest 22*, 3 (2021), 110–161. 2

[Fri08] FRIENDLY M.: A Brief History of Data Visualization. In *Handbook of Data Visualization*, Chen C.-h., Härdle W., Unwin A., (Eds.), Springer Handbooks Comp.Statistics. Springer, Berlin, Heidelberg, 2008, pp. 15–56. doi:10.1007/978-3-540-33037-0_2. 3

[Fry05] FRYSINGER S. P.: A brief history of auditory data representation to the 1980s. In *Proceedings of the International Conference on Auditory Display* (2005). URL: http://hdl.handle.net/1853/50089. 3

[GDAS*18] GUNE A., DE AMICIS R., SIMÕES B., SANCHEZ C. A., DEMIREL H. O.: Graphically hearing: Enhancing understanding of geospatial data through an integrated auditory and visual experience. *IEEE Computer Graphics and Applications 38*, 4 (July 2018), 18–26. doi:10.1109/MCG.2018.042731655. 7, 8, 13, 22

[GGB05] GUTTMAN S. E., GILROY L. A., BLAKE R.: Hearing what the eyes see: auditory encoding of visual temporal sequences. *Psychological science 16*, 3 (2005), 228–235. 4, 22

[GH12] GROND F., HERMANN T.: Singing function: Exploring auditory graphs with a vowel based sonification. *Journal on Multimodal User Interfaces 5*, 3-4 (May 2012), 87–95. doi:10.1007/s12193-011-0068-2. 3

[GHJ*19] GROSS-VOGT K., HERMANN T., JURY M. W., STEINER A. K., KARTADINATA S.: Klima|Anlage—performing climate data. In *Addressing the Challenges in Communicating Climate Change Across Various Audiences*, Leal Filho W., Lackner B., McGhie H., (Eds.). Springer, Cham, 2019, pp. 339–355. doi:10.1007/978-3-319-98294-6_21. 21

[GKW21] GROPPE S., KLINCKENBERG R., WARNKE B.: Sound of databases: Sonification of a semantic web database engine. *Proc. VLDB Endow. 14*, 12 (July 2021), 2695–2698. doi:10.14778/3476311.3476322. 7, 8, 14

[Gou17] GOUDARZI V.: *Systematic Procedure to Develop Sonifications*. PhD thesis, University of Music and Performing Arts Graz, 2017. 17

[GR11] GOMEZ I., RAMIREZ R.: A data sonification approach to cognitive state identification. In *Proc. 17th International Conference on Auditory Display (ICAD-2011)* (Budapest, Hungary, June 2011). http://hdl.handle.net/1853/51569. 7, 8, 13

[GRK*16] GIONFRIDA L., ROGINSKA A., KEARY J., MOHANRAJ H., FRIEDMAN K. P.: The triple tone sonification method to enhance the diagnosis of alzheimer's dementia. In *Proc. 22nd International Conference on Auditory Display (ICAD–2016)* (Canberra, Australia, July 2016). doi:10.21785/icad2016.023. 7, 8, 14, 16, 22

[GS90] GRINSTEIN G. G., SMITH S.: The perceptualization of scientific data. In *Extracting Meaning from Complex Data: Processing, Display, Interaction* (1990), vol. 1259, SPIE, pp. 190–199. doi:10.1117/12.19986. 4

[GVEZ23] GROSS-VOGT K., ENGE K., ZMÖLNIG I.: Reflecting on qualitative and quantitative data to frame criteria for effective sonification design. In *Proceedings of the 18th International Audio Mostly Conference* (New York, NY, USA, 2023), AM '23, Association for Computing Machinery, p. 93–100. doi:10.1145/3616195.3616233. 6

[HAR16] HILDEBRANDT T., AMERBAUER F., RINDERLE-MA S.: Combining sonification and visualization for the analysis of process execution data. In *2016 IEEE 18th Conference on Business Informatics (CBI)* (Aug. 2016), vol. 02, pp. 32–37. doi:10.1109/CBI.2016.47. 7, 8, 12, 13, 14, 17

[HAS11] HULLMAN J., ADAR E., SHAH P.: Benefitting InfoVis with visual difficulties. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (Dec. 2011), 2213–2222. doi:10.1109/TVCG.2011.175. 22

[HCTP14] HOLTZMAN B., CANDLER J., TURK M., PETER D.: Seismic sound lab: Sights, sounds and perception of the earth as an acoustic space. In *Sound, Music, and Motion* (Cham, 2014), Aramaki M., Derrien O., Kronland-Martinet R., Ystad S., (Eds.), Lecture Notes in Computer Science, Springer International Publishing, pp. 161–174. doi:10.1007/978-3-319-12976-1_10. 7, 8, 12

[Her02] HERMANN T.: *Sonification for Exploratory Data Analysis*. PhD thesis, Bielefeld University, Bielefeld, Germany, 02 2002. 10, 16

[Her08] HERMANN T.: Taxonomy and definitions for sonification and auditory display. In *Proceedings of the 14th International Conference on Auditory Display* (2008). URL: http://hdl.handle.net/1853/49960. 3

[Her20] HERRMANN V.: Visualizing and sonifying how an artificial ear hears music. In *Proceedings of the NeurIPS 2019 Competition and Demonstration Track* (Dec. 2020), Escalante H. J., Hadsell R., (Eds.), vol. 123 of *Proceedings of Machine Learning Research*, PMLR, pp. 192–202. https://proceedings.mlr.press/v123/herrmann20a.html. 7, 8, 14

[HEUHEB23] HOQUE M. N., EHTESHAM-UL-HAQUE M., ELMQVIST N., BILLAH S. M.: Accessible data representation with natural sound. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (2023), pp. 826:1–826:19. doi:10.1145/3544548.3581087. 19

[HGI*22] HOLLOWAY L. M., GONCU C., ILSAR A., BUTLER M., MARRIOTT K.: Infosonics: Accessible infographics for people who are blind using sonification and voice. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (2022), pp. 480:1–480:13. doi:10.1145/3491102.3517465. 19

[HH04] HUNT A., HERMANN T.: The importance of interaction in sonification. In *Proc. of the 10th Meeting of the International Conference on Auditory Display (ICAD 2004)* (Sydney, Australia, 2004), pp. 1–8. 19

[HHG*12] HÖFERLIN B., HÖFERLIN M., GOLOUBETS B., HEIDEMANN G., WEISKOPF D.: Auditory support for situation awareness in video surveillance. In *Proceedings of the 18th International Conference on Auditory Display,* (Atlanta, GA, USA,, June 2012), Georgia Institute of Technology, pp. 156–163. URL: http://hdl.handle.net/1853/44426. 20

[HHN11] HERMANN T., HUNT A., NEUHOFF J. G. (Eds.): *The Sonification Handbook*. Logos, Bielefeld, 2011. 2, 3, 19

[HHRM16] HILDEBRANDT T., HERMANN T., RINDERLE-MA S.: Continuous sonification enhances adequacy of interactions in peripheral process monitoring. *International Journal of Human-Computer Studies 95* (2016), 54–65. doi:10.1016/j.ijhcs.2016.06.002. 20

[HK22] HAN Y. C., KHANDUJA A.: The future is red: Visualizing wild-fire predictions using contactless interaction. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2022), CHI EA '22, Association for Computing Machinery. doi:10.1145/3491101.3519903. 7, 8, 10, 13, 16, 17, 22

[HPDW23] HUPPENKOTHEN D., PAMPIN J., DAVENPORT J. R. A., WENLOCK J.: The sonified hertzsprung-russell diagram. In *Proc. 28th International Conference on Auditory Display (ICAD 2023)* (Norrköping, Sweden, June 2023), Georgia Institute of Technology, pp. 272–279. doi:10.21785/icad2023.6263. 7, 8, 10, 12, 16, 17

[HR21] HERMANN T., REINSCH D.: Sc3nb: A Python-SuperCollider Interface for Auditory Data Science. In *Proceedings of the 16th International Audio Mostly Conference* (New York, NY, USA, Oct. 2021), AM '21, Association for Computing Machinery, pp. 208–215. doi:10.1145/3478384.3478401. 22

[HRM15] HILDEBRANDT T., RINDERLE-MA S.: Server sounds and network noises. In *Cognitive Infocommunications (CogInfoCom), 2015 6th IEEE International Conference on* (2015), IEEE, pp. 45–50. 5, 20

[HTHB22] HARRISON C., TRAYFORD J., HARRISON L., BONNE N.: Audio universe: Tour of the solar system. *Astronomy & Geophysics 63*, 2 (Apr. 2022), 2.38–2.40. doi:10.1093/astrogeo/atac027. 7, 8, 16, 17

[Hun07] HUNTER J. D.: Matplotlib: A 2d graphics environment. *Computing in Science & Engineering 9*, 3 (2007), 90–95. doi:10.1109/MCSE.2007.55. 22

[HWKK23] HUANG Z., WITSCHARD D., KUCHER K., KERREN A.: VA + Embeddings STAR: A State-of-the-Art Report on the Use of Embeddings in Visual Analytics. *Computer Graphics Forum 42*, 3 (2023), 539–571. doi:10.1111/cgf.14859. 19

[Ibe20] IBER M.: Auditory display in workplace environments. In *Foundations in Sound Design for Embedded Media: A Multidisciplinary Approach*, Filimowicz M., (Ed.). Routledge, New York, 2020, pp. 131–154. 4, 20

[IIC*13] ISENBERG T., ISENBERG P., CHEN J., SEDLMAIR M., MÖLLER T.: A systematic review on the practice of evaluating visualization. *IEEE Trans. Visualization and Computer Graphics 19*, 12 (2013), 2818–2827. doi:10.1109/TVCG.2013.126. 15

[JBKS15] JAIN A., BANSAL R., KUMAR A., SINGH K.: A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students. *International Journal of Applied and Basic Medical Research 5* (2015), 124. doi:10.4103/2229-516X.157168. 4

[JDS15] JAKUS G., DICKE C., SODNIK J.: A user study of auditory, head-up and multi-modal displays in vehicles. *Applied Ergonomics 46* (2015), 184–192. doi:10.1016/j.apergo.2014.08.008. 20

[JLSP15] JEON M., LEE J.-H., STERKENBURG J., PLUMMER C.: *Cultural Differences in Preference of Auditory Emoticons: USA and South Korea*. Georgia Institute of Technology, July 2015. 2

[JMP13] JOLIAT N., MAYTON B., PARADISO J. A.: Spatialized anonymous audio for browsing sensor networks via virtual worlds. In *Proc. 19th International Conference on Auditory Display (ICAD2013)* (Lodz, Poland, July 2013), Georgia Institute of Technology, pp. 67–75. http://hdl.handle.net/1853/51643. 7, 8, 16

[Joh04] JOHANNSEN G.: Auditory displays in human-machine interfaces. *Proceedings of the IEEE 92*, 4 (2004), 742–758. doi:10.1109/JPROC.2004.825905. 20

[KAV21] KARIYADO Y., AREVALO C., VILLEGAS J.: Auralization of three-dimensional cellular automata. In *Artificial Intelligence in Music, Sound, Art and Design* (Cham, 2021), Romero J., Martins T., Rodríguez-Fernández N., (Eds.), Lecture Notes in Computer Science, Springer International Publishing, pp. 161–170. doi:10.1007/978-3-030-72914-1_11. 7, 8, 15, 22

[KBBG07] KASAKEVICH M., BOULANGER P., BISCHOF W. F., GARCIA

M.: Augmentation of visualisation using sonification: A case study in computational fluid dynamics. In *Proc. IPT-EGVE Symposium* (Germany, Europe, 2007), The Eurographics Association, pp. 89–94. doi:10.2312/PE/VE2007Short/089-094. 4

[KGS*20] KIM C., GUO A., SALHOTRA G., SPRINKHUIZEN S., SHETTY K., KONG D. S.: Sonifying data from the human microbiota: Biota beats. *Computer Music Journal 44*, 1 (Mar. 2020), 51–70. doi:10.1162/comj_a_00552. 21

[KIK19] KANEV I. K., ILIEV I. T., KRASTEVA V. T.: Sonification – an alternative presentation of the electrocardiogram: A systematic literature review. In *2019 IEEE XXVIII International Scientific Conference Electronics (ET)* (2019), IEEE, pp. 1–4. doi:10.1109/ET.2019.8878650. 5, 20

[KJRK21] KIM N. W., JOYNER S. C., RIEGELHUTH A., KIM Y.-S.: Accessible visualization: Design space, opportunities, and challenges. *Computer Graphics Forum 40*, 3 (2021), 173–188. doi:10.1111/cgf.14298. 19

[KLTW17] KONDAK Z., LIANG T. A., TOMLINSON B., WALKER B. N.: Web sonification sandbox - an easy-to-use web application for sonifying data and equations. In *Proc. Web Audio Conference WAC-2017* (London, UK, Aug. 2017). https://qmro.qmul.ac.uk/xmlui/handle/123456789/26083. 7, 8, 11, 14, 22

[KOR23] KIESENHOFER S., OELSCH A., RAMMER D.: Deep sync, 2023. Webpage: https://ars.electronica.art/futurelab/en/projects-deep-sync/. Last accessed: Dec 28st 2023. 21

[Kra94] KRAMER G. (Ed.): *Auditory Display: Sonification, Audification and Auditory Interfaces. Proceedings of the First International Conference on Auditory Display (ICAD) 1992*. Addison-Wesley, Reading, Mass, 1994. 3

[KWB*99] KRAMER G., WALKER B., BONEBRIGHT T., COOK P., FLOWERS J. H., MINER N., NEUHOFF J., ET AL.: *Sonification Report: Status of the Field and Research Agenda*. Report for the NSF, International Community for Auditory Display, 1999. URL: http://www.icad.org/websiteV2.0/References/nsf.html. 1, 3

[LCS*28] LIN J.-W., CHEN W., SHEN C.-P., CHIU M.-J., KAO Y.-H., LAI F., ZHAO Q., CICHOCKI A.: Visualization and sonification of long-term epilepsy electroencephalogram monitoring. *Journal of Medical and Biological Engineering 38*, 6 (2028), 943–952. doi:10.1007/s40846-017-0358-6. 20

[LDIC21] LEE B., DACHSELT R., ISENBERG P., CHOE E. K. (Eds.): *Mobile Data Visualization*. Chapman and Hall/CRC, New York, Dec. 2021. doi:10.1201/9781003090823. 21

[Len21] LENZI S.: *The Design of Data Sonification. Design Processes, Protocols and Tools Grounded in Anomaly Detection*. PhD thesis, Politecnico di Milano, 2021. https://www.politesi.polimi.it/handle/10589/177079. 3

[LF21] LINDETORP H., FALKENBERG K.: Sonification for everyone everywhere: Evaluating the webaudioxml sonification toolkit for browsers. In *Proc. 26th International Conference on Auditory Display (ICAD 2021)* (Virtual Conference, June 2021), Georgia Institute of Technology, pp. 15–21. http://hdl.handle.net/1853/66351. 7, 8, 11, 12, 14, 22

[LLC23] LINDBORG P., LENZI S., CHEN M.: Climate data sonification and visualization: An analysis of topics, aesthetics, and characteristics in 32 recent projects. *Frontiers in Psychology 13* (2023). doi:10.3389/fpsyg.2022.1020102. 20

[LLW21] LYU Z., LI J., WANG B.: Aiive: Interactive visualization and sonification of neural networks in virtual reality. In *2021 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)* (Nov. 2021), pp. 251–255. doi:10.1109/AIVR52153.2021.00057. 7, 8, 10

[LSB*23] LEMMON E., SCHEDEL M., BILKHU I., ZHU H., ESCOBAR L., AUMOITHE G.: Mapping in the emergency: Designing a hyperlocal and socially conscious sonified map of Covid-19 in Suffolk County, New York. In *Proceedings of ISon 2022, 7th Interactive Sonification Workshop,*

BSCC, University of Bremen (Jan. 2023), Zenodo. `doi:10.5281/ZENODO.7552257`. 7, 8, 10, 14, 16

[LWP99] LOTTO R. B., WILLIAMS S. M., PURVES D.: Mach bands as empirically derived associations. *Proceedings of the National Academy of Sciences 96*, 9 (1999), 5245–5250. 4

[LWP*13] LU S. A., WICKENS C. D., PRINET J. C., HUTCHINS S. D., SARTER N., SEBOK A.: Supporting interruption management and multimodal interface design: Three meta-analyses of task performance as a function of interrupting task modality. *Human Factors 55*, 4 (2013), 697–724. `doi:10.1177/0018720813476298`. 20

[LWS96] LODHA S. K., WILSON C. M., SHEEHAN R. E.: Listen: Sounding uncertainty visualization. In *Visualization '96. Proceedings.* (San Francisco, CA, USA, 1996), ACM, pp. 189–195. `doi:10.1109/VISUAL.1996.568105`. 4

[MAFP19a] MALIKOVA E., ADZHIEV V., FRYAZINOV O., PASKO A.: Visual-auditory representation and analysis of molecular scalar fields. In *Proc. EUROGRAPHICS 2019* (2019), pp. 25–26. `doi:10.2312/egp.20191051`. 17

[MAFP19b] MALIKOVA E., ADZHIEV V., FRYAZINOV O., PASKO A.: Visual-auditory volume rendering of scalar fields. In *Proc. 25th International Conference on Auditory Display (ICAD 2019)* (Northumbria University, June 2019), Georgia Institute of Technology, pp. 147–154. `doi:10.21785/icad2019.004`. 6, 7, 8, 13

[May14] MAYER R. E.: The cambridge handbook of multimedia learning. *The Cambridge Handbook of Multimedia Learning, Second Edition* (1 2014), 1–930. `doi:10.1017/CBO9781139547369`. 4

[MBKSM16] METATLA O., BRYAN-KINNS N., STOCKMAN T., MARTIN F.: Sonification of reference markers for auditory graphs: Effects on non-visual point estimation tasks. *PeerJ Computer Science 2* (2016), e51. 22

[MF95] MINGHIM R., FORREST A.: An illustrated analysis of sonification for scientific visualisation. In *Proceedings Visualization '95* (1995), pp. 110–117. `doi:10.1109/VISUAL.1995.480802`. 3, 4

[ML17] MCNABB L., LARAMEE R. S.: Survey of Surveys (SoS) - Mapping The Landscape of Survey Papers in Information Visualization. *Computer Graphics Forum 36*, 3 (2017), 589–617. `doi:10.1111/cgf.13212`. 4

[MLS95] MOUSAVI S. Y., LOW R., SWELLER J.: Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of educational psychology 87*, 2 (1995), 319. 19

[MMM18] MAÇÃS C., MARTINS P., MACHADO P.: Consumption as a rhythm: A multimodal experiment on the representation of time-series. In *2018 22nd International Conference Information Visualisation (IV)* (July 2018), pp. 504–509. `doi:10.1109/iV.2018.00093`. 7, 8, 13, 14

[MMU16] MATSUBARA M., MORIMOTO Y., UCHIDE T.: Collaborative study of interactive seismic array sonification for data exploration and public outreach activities. In *Proceedings of ISon 2016, 5th Interactive Sonification Workshop, CITEC, Bielefeld University* (Bielefeld, Germany, 2016). 7, 8, 12, 14, 16

[MNS16] MACULEWICZ J., NILSSON N. C., SERAFIN S.: An investigation of the effect of immersive visual and auditory feedback on rhythmic walking interaction. In *Proceedings of the Audio Mostly 2016* (2016), AM '16, Association for Computing Machinery, pp. 194–201. `doi:10.1145/2986416.2986429`. 20

[MNW*18] MACDONALD D. E., NATARAJAN T., WINDEYER R. C., COPPIN P., STEINMAN D. A.: Data-driven sonification of cfd aneurysm models. In *Proc. 24th International Conference on Auditory Display (ICAD 2018)* (Houghton, Michigan, June 2018), The International Community for Auditory Display, pp. 28–33. `doi:10.21785/icad2018.010`. 7, 8, 15, 16, 22

[Mun15] MUNZNER T.: *Visualization Analysis and Design*. CRC Press, Boca Raton, Jan. 2015. 1, 6, 10, 11, 12, 13, 14, 17, 22

[MVC*10] MASUD L., VALSECCHI F., CIUCCARELLI P., RICCI D., CAVIGLIA G.: From data to knowledge – visualizations as transformation processes within the data-information-knowledge continuum. In *2010 14th International Conference Information Visualisation* (July 2010), pp. 445–449. `doi:10.1109/IV.2010.68`. 21

[NB02] NESBITT K. V., BARRASS S.: Evaluation of a multimodal sonification and visualisation of depth of market stock data. In *Proceedings of the 8th International Conference on Auditory Display* (Kyoto, 2002). URL: http://hdl.handle.net/1853/51355. 4

[Nee19] NEES M. A.: Eight components of a design theory of sonification. In *Proceedings of the 25th International Conference on Auditory Display (ICAD 2019)* (2019), pp. 176–183. `doi:10.21785/icad2019.048`. 3

[Neu04] NEUHOFF J. G. (Ed.): *Ecological Psychoacoustics*. Elsevier Academic Press, San Diego, 2004. 3, 22

[Neu19] NEUHOFF J. G.: Is sonification doomed to fail. In *Proceedings of the 25th International Conference on Auditory Display (ICAD 2019)* (2019), pp. 327–330. 3

[NGJW14] NEES M. A., GABLE T. M., JEON M., WALKER B. N.: Prototype auditory displays for a fuel efficiency driver interface. In *Proceedings of the 20th International Conference on Auditory Display* (New York, USA, June 2014), Georgia Institute of Technology. Publisher: Georgia Institute of Technology. URL: http://hdl.handle.net/1853/52089. 20

[NRL*12] NESS S., REIMER P., LOVE J., SCHLOSS W. A., TZANETAKIS G.: Sonophenology. *Journal on Multimodal User Interfaces 5*, 3 (May 2012), 123–129. `doi:10.1007/s12193-011-0066-4`. 7, 8, 14, 16

[NSC16] NORTH K. J., SARMA A., COHEN M. B.: Understanding git history: A multi-sense view. In *Proceedings of the 8th International Workshop on Social Software Engineering* (New York, NY, USA, 2016), SSE 2016, Association for Computing Machinery, pp. 1–7. `doi:10.1145/2993283.2993285`. 7, 8, 12, 14, 17, 22

[NW08] NEES M. A., WALKER B. N.: Encoding and representation of information in auditory graphs: Descriptive reports of listener strategies for understanding data. In *Proceedings of the International Conference on Auditory Display* (Paris, FR, June 2008). URL: http://hdl.handle.net/1853/49907. 3

[PBM15] PAPACHRISTODOULOU P., BETELLA A., MANZOLLI J.: Augmenting the navigation of complex data sets using sonification: A case study with brainx3. In *2015 IEEE 2nd VR Workshop on Sonic Interactions for Virtual Environments (SIVE)* (Mar. 2015), pp. 1–6. `doi:10.1109/SIVE.2015.7361284`. 7, 8, 12, 16, 22

[PBV14] PAPACHRISTODOULOU P., BETELLA A., VERSCHURE P.: Sonification of large datasets in a 3d immersive environment: A neuroscience case study. In *Proc. ACHI 2014 : The Seventh International Conference on Advances in Computer-Human Interactions* (Jan. 2014), IARIA, pp. 35–40. 7, 8, 22

[PC19] PHILLIPS S., CABRERA A.: Sonification workstation. In *25th International Conference on Auditory Display (ICAD 2019)* (Northumbria University, June 2019), Georgia Institute of Technology, pp. 184–190. http://hdl.handle.net/1853/61529. 7, 8, 11, 12, 14, 22

[PCB23] PENG T., CHOI H., BERGER J.: Siren: Creative and extensible sonification on the web. In *Proc. 28th International Conference on Auditory Display (ICAD 2023)* (Norrköping, Sweden, June 2023), Georgia Institute of Technology, pp. 78–84. https://hdl.handle.net/1853/72879. 7, 8, 11, 14, 22

[PFH*22] PATÉ A., FARGE G., HOLTZMAN B. K., BARTH A. C., POLI P., BOSCHI L., KARLSTROM L.: Combining audio and visual displays to highlight temporal and spatial seismic patterns. *Journal on Multimodal User Interfaces 16*, 1 (Mar. 2022), 125–142. `doi:10.1007/s12193-021-00378-8`. 7, 8, 10, 12, 14, 16, 22

[Pol06] POLLI A.: Heat and the Heartbeat of the City: Sonifying Data Describing Climate Change. *Leonardo Music Journal 16* (Dec. 2006), 44–45. `doi:10.1162/lmj.2006.16.44`. 21

[RC15] ROSLI M. H. W., CABRERA A.: Gestalt principles in multimodal

data representation. *IEEE Computer Graphics & Applications 32* (2015), 80–87. doi:10.1109/MCG.2015.29. 4

[RCL*98] ROBERTSON G., CZERWINSKI M., LARSON K., ROBBINS D. C., THIEL D., VAN DANTZICH M.: Data mountain: Using spatial memory for document management. In *Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 1998), UIST '98, ACM, pp. 153–162. doi:10.1145/288392.288596. 4

[RF22] RÖNNBERG N., FORSELL C.: Questionnaires assessing usability of audio-visual representations. In *Proc. AVI 2022 Workshop on Audio-Visual Analytics* (Rome, 2022). doi:10.5281/zenodo.6555676. 15, 21

[RFK*15] RAU B., FRIESS F., KRONE M., MULLER C., ERTL T.: Enhancing visualization of molecular simulations using sonification. In *2015 IEEE 1st International Workshop on Virtual and Augmented Reality for Molecular Science (VARMS@IEEEVR)* (Mar. 2015), pp. 25–30. doi:10.1109/VARMS.2015.7151725. 7, 8, 12, 13, 15

[RFM13] ROGIŃSKA A., FRIEDMAN K., MOHANRAJ H.: Exploring sonification for augmenting brain scan data. In *Proc. 19th International Conference on Auditory Display (ICAD2013)* (Lodz, Poland, July 2013), Georgia Institute of Technology, pp. 95–101. http://hdl.handle.net/1853/51653. 7, 8

[RG71] RUBINSTEIN L., GRUENBERG E. M.: Intramodal and crossmodal sensory transfer of visual and auditory temporal patterns. *Perception & Psychophysics 9* (9 1971), 385–390. doi:10.3758/BF03210235. 4

[RH22] REINSCH D., HERMANN T.: Interacting with sonifications: The mesonic framework for interactive auditory data science. In *Proceedings of the 7th Interactive Sonification Workshop (ISon)* (2022), CITEC, Bielefeld University, p. 65. 22

[RH23] REINSCH D., HERMANN T.: Sonecules: A Python sonification architecture. In *Proceedings of the 28th International Conference on Auditory Display* (Norrköping, Sweden, June 2023), Georgia Institute of Technology, pp. 62–69. doi:10.21785/icad2023.5580. 22

[Rib18] RIBER A. G.: Planethesizer: Approaching exoplanet sonification. In *Proc. 24th International Conference on AuditoryDisplay (ICAD 2018)* (Michigan Technological University, June 2018), Georgia Institute of Technology, pp. 219–226. http://hdl.handle.net/1853/60073. 7, 8, 14

[Rib19] RIBER A. G.: Sonifigrapher: Sonified light curve synthesizer. In *Proc. 25th International Conference on Auditory Display (ICAD 2019)* (Northumbria University, June 2019), Georgia Institute of Technology, pp. 62–66. http://hdl.handle.net/1853/61497. 7, 8, 11, 12, 13, 14, 16

[RJ16] RÖNNBERG N., JOHANSSON J.: Interactive sonification for visual dense data displays. In *Proc. 5th Interactive Sonification Workshop, ISon* (2016), CITEC, Bielefeld University, pp. 63–67. 7, 8, 13, 14, 15, 22

[RMS*22] ROCHE T. R., MAAS E. J. C., SAID S., BRAUN J., MACHADO C., SPAHN D. R., NOETHIGER C. B., TSCHOLL D. W.: Anesthesia personnel's visual attention regarding patient monitoring in simulated non-critical and critical situations, an eye-tracking study. *BMC Anesthesiology 22*, 1 (2022), 167:1–167:10. doi:10.1186/s12871-022-01705-6. 20

[RN96] ROSSITER D., NG W.-Y.: A system for the complementary visualization of 3d volume images using 2d and 3d binaurally processed sonification representations. In *Proceedings of Seventh Annual IEEE Visualization '96* (1996), pp. 351–354. doi:10.1109/VISUAL.1996.568129. 4

[Rön19a] RÖNNBERG N.: Musical sonification supports visual discrimination of color intensity. *Behaviour & Information Technology 38*, 10 (2019), 1028–1037. doi:10.1080/0144929X.2019.1657952. 19, 22

[Rön19b] RÖNNBERG N.: Sonification supports perception of brightness contrast. *Journal on Multimodal User Interfaces 13*, 4 (2019), 373–381. doi:10.1007/s12193-019-00311-0. 19, 22

[Rön21] RÖNNBERG N.: Sonification for conveying data and emotion. In *Proceedings of the 16th International Audio Mostly Conference* (New York, NY, USA, 2021), AM '21, Association for Computing Machinery, pp. 56–63. doi:10.1145/3478384.3478387. 7, 8, 10, 14, 22

[RS22] RUSSO M., SANTAGUIDA A.: 5000 exoplanets: Listen to the sounds of discovery. In *Proc. 27th International Conference on Auditory Display (ICAD 2022)* (Virtual Conference, June 2022), Georgia Institute of Technology, pp. 64–68. http://hdl.handle.net/1853/67384. 7, 8, 10, 13, 14, 16, 17, 22

[RW10] ROBERTS J. C., WALKER R.: Using all our senses: The need for a unified theoretical approach to multi-sensory information visualization. In *Workshop on the Role of Theory in Information Visualization* (2010). 3

[RYTW23] RUBAB S., YU L., TANG J., WU Y.: Exploring Effective Relationships Between Visual-Audio Channels in Data Visualization. *Journal of Visualization 26*, 4 (Aug. 2023), 937–956. doi:10.1007/s12650-023-00909-3. 4

[SAR22] SVORONOS-KANAVAS I., AGIOMYRGIANAKIS V., RÖNNBERG N.: An exploratory use of audiovisual displays on oceanographic data. In *Proc. AVI 2022 Workshop on Audio-Visual Analytics (WAVA22)* (Frascati, Italy, 2022), Zenodo. doi:10.5281/zenodo.6555839. 7, 8, 13

[SBG90] SMITH S., BERGERON R. D., GRINSTEIN G. G.: Stereophonic and surface sound generation for exploratory data analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 1990), CHI '90, Association for Computing Machinery, pp. 125–132. doi:10.1145/97243.97264. 4

[SDS23] SPENCE C., DI STEFANO N.: Sensory translation between audition and vision. *Psychonomic Bulletin & Review* (Oct. 2023). doi:10.3758/s13423-023-02343-w. 4

[SFLD22] SIRKKA A., FAGERLÖNN J., LINDBERG S., DELSING K.: The design of an auditory alarm concept for a paper mill control room. In *Advances in Ergonomics In Design, Usability & Special Populations: Part III* (2022), vol. 20, AHFE Open Acces. ISSN: 27710718 Issue: 20. doi:10.54941/ahfe1001301. 20

[SGS*18] SCHLOSS K. B., GRAMAZIO C. C., SILVERMAN A. T., PARKER M. L., WANG A. S.: Mapping color to meaning in colormap data visualizations. *IEEE transactions on visualization and computer graphics 25*, 1 (2018), 810–819. 4

[Shn96] SHNEIDERMAN B.: The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the IEEE Symposium on Visual Languages, VL* (1996), pp. 336–343. doi:10.1109/VL.1996.545307. 3

[SJMT19] SCHAFFERT N., JANZEN T. B., MATTES K., THAUT M. H.: A review on the relationship between sound and movement in sports and rehabilitation. *Frontiers in Psychology 10* (2019). doi:10.3389/fpsyg.2019.00244. 5, 20

[SKS06] SEITZ A. R., KIM R., SHAMS L.: Sound facilitates visual learning. *Current biology : CB 16* (7 2006), 1422–1427. doi:10.1016/J.CUB.2006.05.048. 4

[SMM12] SEDLMAIR M., MEYER M., MUNZNER T.: Design study methodology: Reflections from the trenches and the stacks. *IEEE Trans. Visualization and Computer Graphics 18*, 12 (2012), 2431–2440. doi:10.1109/TVCG.2012.213. 17

[SMS21] SCHLOSSER P. D., MATTHEWS B., SANDERSON P. M.: Head-worn displays for healthcare and industry workers: A review of applications and design. *International Journal of Human-Computer Studies 154* (2021), 102628. doi:10.1016/j.ijhcs.2021.102628. 20

[SNH05] STOCKMAN T., NICKERSON L. V., HIND G.: Auditory graphs: A summary of current experience and towards a research agenda. In *International Conference on Auditory Display (ICAD2005), Limerick, Ireland* (2005). 19

[SNHS13] SCHULZ H.-J., NOCKE T., HEITZLER M., SCHUMANN H.: A design space of visualization tasks. *IEEE Trans. Visualization and Computer Graphics 19*, 12 (Dec. 2013), 2366–2375. doi:10.1109/TVCG.2013.120. 3

[SPA23] STOIBER C., POHL M., AIGNER W.: Design actions for the design of visualization onboarding methods. In *EduVis Workshop, IEEE VIS* (Melbourne, Australia, 2023). doi:10.31219/osf.io/wjp5x. 23

[Spe07] SPENCE R.: *Information Visualization: Design for Interaction*, 2nd ed. Prentice Hall, Upper Saddle River, NJ, 2007. 3

[Spe11] SPENCE C.: Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics 2011 73:4 73* (1 2011), 971–995. `doi:10.3758/S13414-010-0073-7`. 4

[SQS*20] SU I., QIN Z., SARACENO T., BISSHOP A., MÜHLETHALER R., ZIPORYN E., BUEHLER M. J.: Sonification of a 3-D spider web and reconstitution for musical composition using granular synthesis. *Computer Music Journal 44*, 4 (Dec. 2020), 43–59. `doi:10.1162/comj_a_00580`. 21

[SŠ18] SEBJANIČ R., ŠUTIĆ G.: Aqua_forensic, 2018. Webpage: https://ars.electronica.art/aeblog/en/2018/11/07/aquaforensic/. Last accessed: Dec 21st 2023. 21

[Sup12] SUPPER A.: *Lobbying for the Ear : The Public Fascination with and Academic Legitimacy of the Sonification of Scientific Data*. Doctoral thesis, Maastricht University, 2012. `doi:10.26481/dis.20120606as`. 3

[SW13] SCHUETT J. H., WALKER B. N.: Measuring comprehension in sonification tasks that have multiple data streams. In *Proc. of the Audio Mostly Conference* (2013), pp. 11:1–11:6. `doi:10.1145/2544114.2544121`. 4

[SWG*23] STOIBER C., WAGNER M., GRASSINGER F., POHL M., STITZ H., STREIT M., POTZMANN B., AIGNER W.: Visualization onboarding grounded in educational theories. In *Visualization Psychology*. Springer Nature, 2023. `doi:10.1007/978-3-031-34738-2`. 23

[TB23] TRAVER P., BERGH E.: Harmonices solaris - sonification of the planets. In *Proc. 28th International Conference on Auditory Display (ICAD 2023)* (Norrköping, Sweden, June 2023), Georgia Institute of Technology, pp. 242–248. `doi:10.21785/icad2023.204`. 7, 8, 17

[TC05] THOMAS J. J., COOK K. A. (Eds.): *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE, Los Alamitos, CA, USA, 2005. 4

[TML*15] TARDIEU J., MISDARIIS N., LANGLOIS S., GAILLARD P., LEMERCIER C.: Sonification of in-vehicle interface reduces gaze movements under dual-task condition. *Applied Ergonomics 50* (2015), 41–49. `doi:10.1016/j.apergo.2015.02.004`. 20

[TMN*21] TEMOR L., MACDONALD D. E., NATARAJAN T., COPPIN P. W., STEINMAN D. A.: Perceptually-motivated sonification of spatiotemporally-dynamic CFD data. In *Proc. 26th International Conference on Auditory Display (ICAD 2021)* (Virtual Conference, June 2021), Georgia Institute of Technology, pp. 202–209. http://hdl.handle.net/1853/66343. 7, 8, 15, 16, 22

[VG97] VOLPE C. R., GLINERT E. P.: Auralization of streamline vorticity in computational fluid dynamics data. In *Visualization '97., Proceedings* (1997), pp. 51–57. `doi:10.1109/VISUAL.1997.663856`. 4

[VH06] VICKERS P., HOGG B.: Sonification abstraite/sonification concrète: An 'æsthetic persepctive space' for classifying auditory displays in the ars musica domain. In *Proceedings of the 12th International Conference on Auditory Display* (London, UK, June 2006), Georgia Institute of Technology, pp. 210–216. URL: http://hdl.handle.net/1853/50641. 3

[Vic16] VICKERS P.: Sonification and music, music and sonification. In *The Routledge Companion to Sounding Art*, Cobussen M., Meelberg V., Truax B., (Eds.). Routledge, 2016, pp. 135–144. 10

[VRGM20] VAN RHEDEN V., GRAH T., MESCHTSCHERJAKOV A.: Sonification approaches in sports in the past decade: a literature review. In *Proceedings of the 15th International Audio Mostly Conference* (2020), ACM, pp. 199–205. `doi:10.1145/3411109.3411126`. 5, 20

[War19] WARE C.: *Information visualization: perception for design*. Morgan Kaufmann, 2019. 4

[WHT06] WARD J., HUCKSTEP B., TSAKANIKOS E.: Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex 42* (2006), 264–280. `doi:10.1016/S0010-9452(08)70352-6`. 4

[Wil05] WILKINSON L.: *The Grammar of Graphics*, second ed. Springer, New York, 2005. 3

[WL01] WALKER B. N., LANE D. M.: Sonification mappings database on the web. In *Proceedings of the 2001 International Conference on Auditory Display* (Espoo, Finland, 2001), Georgia Institute of Technology, p. 281. 4

[WM10] WALKER B. N., MAUNEY L. M.: Universal design of auditory graphs: A comparison of sonification mappings for visually impaired and sighted listeners. *ACM Transactions on Accessible Computing (TACCESS) 2*, 3 (2010), 1–16. 2, 19, 23

[WMY*17] WANG Q., MARKOPOULOS P., YU B., CHEN W., TIMMERMANS A.: Interactive wearable systems for upper body rehabilitation: a systematic review. *Journal of NeuroEngineering and Rehabilitation 14*, 1 (2017), 20. `doi:10.1186/s12984-017-0229-y`. 5, 20

[Woh14] WOHLIN C.: Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering* (London England United Kingdom, May 2014), ACM, pp. 1–10. `doi:10.1145/2601248.2601268`. 6

[Wor19a] WORRALL D.: Polymedia design for network metadata monitoring. In *Sonification Design: From Data to Intelligible Soundfields*, Worrall D., (Ed.). Springer, Cham, 2019, pp. 253–273. `doi:10.1007/978-3-030-01497-1_9`. 20

[Wor19b] WORRALL D.: *Sonification Design: From Data to Intelligible Soundfields*. Springer, Cham, 2019. `doi:10.1007/978-3-030-01497-1`. 3

[WRK*19] WALKER B. N., REHG J. M., KALRA A., WINTERS R. M., DREWS P., DASCALU J., DAVID E. O., DASCALU A.: Dermoscopy diagnosis of cancerous lesions utilizing dual deep learning algorithms via visual and audio (sonification) outputs: Laboratory and prospective observational studies. *EBioMedicine 40* (2019), 176–183. `doi:10.1016/j.ebiom.2019.01.028`. 20

[WS07] WATSON M., SANDERSON P.: Designing for attention with sound: Challenges and extensions to ecological interface design. *Human Factors 49*, 2 (2007), 331–346. `doi:10.1518/001872007X312531`. 20

[WW15] WINTERS R. M., WEINBERG G.: Sonification of the tohoku earthquake: Music, popularization & the auditory sublime. In *Proc. The 21th International Conference on Auditory Display (ICAD–2015* (July 2015), Georgia Institute of Technology, pp. 273–280. http://hdl.handle.net/1853/54149. 7, 8, 11, 12, 13, 14

[XWX*22] XIN X., WANG Y., XIANG G., YANG W., LIU W.: Effectiveness of multimodal display in navigation situation. In *The Ninth International Symposium of Chinese CHI* (2022), Chinese CHI 2021, Association for Computing Machinery, pp. 50–62. `doi:10.1145/3490355.3490361`. 20

[YH17] YANG J., HERMANN T.: Mode explorer: Using model-based sonification to investigate basins of attraction. In *Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences* (New York, NY, USA, 2017), AM '17, Association for Computing Machinery. `doi:10.1145/3123514.3123525`. 8, 12, 13

[YH18] YANG J., HERMANN T.: Interactive mode explorer sonification enhances exploratory cluster analysis. *Journal of the Audio Engineering Society 66*, 9 (Sept. 2018), 703–711. `doi:10.17743/jaes.2018.0042`. 6, 7, 10, 16, 22

[YKSJ07] YI J. S., KANG Y. A., STASKO J. T., JACKO J. A.: Toward a deeper understanding of the role of interaction in information visualization. *IEEE Trans. Visualization and Computer Graphics 13*, 6 (2007), 1224–1231. `doi:10.1109/TVCG.2007.70515`. 3

[Ylm23] YLMAZ S.: Orbuculum, 2023. Ars Electronica Festival 2023. Webpage: https://ars.electronica.art/who-owns-the-truth/de/orbuculum/. Last accessed: Dec 28th 2023. 21

[YM22] YESHCHENKO A., MENDLING J.: A survey of approaches for

event sequence analysis and visualization using the ESeVis framework, 2022. `arXiv:2202.07941`. 6

[Zie23] ZIEMER T.: Three-dimensional sonification for image-guided surgery. In *Proceedings of the 28th International Conference on Auditory Display* (Norrköping, Sweden, June 2023), Georgia Institute of Technology, pp. 8–14. `doi:10.21785/icad2023.2324`. 20

[ZPR16] ZAGERMANN J., PFEIL U., REITERER H.: Measuring cognitive load using eye tracking technology in visual computing. In *Proceedings of the sixth workshop on beyond time and errors on novel evaluation methods for visualization* (2016), ACM, pp. 78–85. `doi:10.1145/2993901.2993908`. 19

[ZTSS23] ZIMNICKI C., TSENG C., SZAFIR D. A., SCHLOSS K. B.: Effects of data distribution and granularity on color semantics for colormap data visualizations. In *2023 IEEE Visualization and Visual Analytics (VIS)* (2023), pp. 11–15. `doi:10.1109/VIS54172.2023.00011`. 4