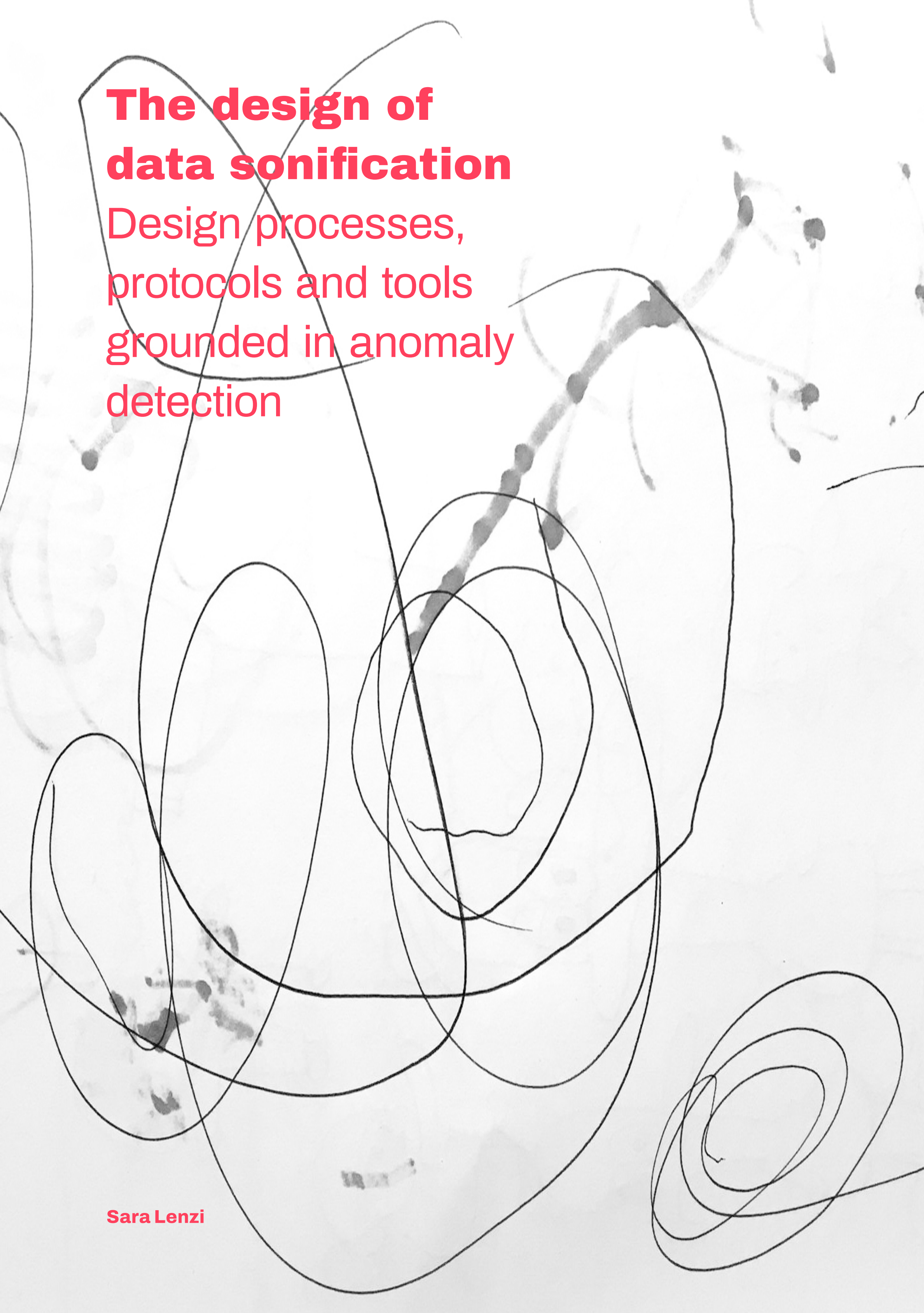
The design of data sonification

Design processes, protocols and tools grounded in anomaly detection



Sara Lenzi

The design of data sonification
Design processes, protocols and tools grounded in anomaly detection

PhD candidate
Sara Lenzi

Supervisors
Paolo Ciuccarelli
Michele Mauri
Juan Sádaba

External examiners
Myounghoon Jeon
Dietmar Offenhuber

PhD Program Coordinator
Paola Bertola

Politecnico di Milano
Design Department
Design Research PhD Program
XXXIII cycle
Milan, June 2021

DOTTORATO
DI RICERCA
IN DESIGN

POLITECNICO
MILANO 1863

Ai miei genitori, Ivo e Lucia.
A Juan.

“Ainsi mon dessein n’est pas d’enseigner ici la méthode que chacun doit suivre pour bien conduire sa raison, mais seulement de fair voir en quelle sorte j’ai tâché de conduire la mienne.”

— *René Descartes.*
Discours de la méthode. Ed. GF - Flammarion. p 17.

Abstract

[EN]

Data sonification is the use of sound to represent and communicate data. It is a field that emerged almost three decades ago, and today, in an increasingly data-intense society, it is gaining momentum as an alternative or complement to the visualization of data. Nonetheless, a series of unresolved issues are still preventing sonification's full transition from a niche practice for the analysis of scientific data to a widely adopted medium that could impact the way we make sense of complex phenomena (Vickers and Barrass 2011; Barrass 2012; Nees 2019). The lack of widely adopted design tools, experimental protocols and shared processes for the design of sonifications, are among the most cited obstacles. As a structured discipline that is used to tackle complex, real-world problems, can design provide the necessary means for data sonification to expand its reach and emerge as a valid alternative for building and shaping our relationship with data?

In this work, I attempt to answer this overarching question in two ways. Through a theoretical investigation, I will first inquire into other sound design practices and their contribution in terms of processes, tools and evaluation protocols; by looking at recent cases, I will later explore the role of intentionality as a pre-requisite for a designerly approach to data sonification. To contextualize intentionality – i.e., deliberate decisions taken to address specific needs, with a purpose and in a given context -I will also engage authors of recent sonifications in a conversation where their approach to sonification will emerge through their projects. From these conversations, I will chart a data sonification space map which is condensed into decisional blocks to form the basis for a data sonification canvas. The canvas, which I intend as a tool to support communication designers in the use of data sonification, will be validated through two series of workshops to provide the first tangible results from this work.

Two design actions form the basis of the practical part and ground the theoretical investigation in data sonification applied to

real-world cases: the detection of anomalies caused by cyber-attacks on digital and digital-physical networks. The two actions involved the design of a series of prototypes through a purposely defined experimental protocol in which expert users were engaged in quantitative and qualitative research in a real-world setting. The definition of a structured design process and the experimental protocol which emerged from these design actions represent the two main results of this work. Lastly, this work contributes to the corpus of experimental evidence on the potentialities of data sonification as a valid solution for the real-time monitoring of anomalies. In particular, the prototype designed for the second design action is now a fully functioning anomaly-detection tool for which development is ongoing.

This work proposes a designerly approach to data sonification which aims to increase the role of sound as a sensory modality for making sense of an increasingly complex world through: a proposal for a structured design process and an experimental protocol - both of which are grounded in sonification for anomaly detection; the definition of a sonification design canvas as a tool to guide designers in the integration of sound as a data representation modality; the development of specific experimentally validated tools that use sonification to support experts in monitoring anomalies in real-world contexts.

[IT]

Quasi trent'anni fa, la sonificazione dei dati nasceva come campo di ricerca specifico, con l'obiettivo di rappresentare e comunicare dati (in quel momento, principalmente numerici) attraverso il suono. Oggi, in una società sempre più dipendente e influenzata dall'uso intensivo di dati, la sonificazione sembra avere l'opportunità di affermarsi come strumento alternativo o complementare alla visualizazzione, non soltanto nel campo dell'analisi scientifica ma anche come strumento di supporto al processo decisionale o come mezzo di comunicazione ad un pubblico non specialistico.

Tuttavia, la sonificazione in quanto campo di ricerca sembra ancora scontare una serie di questioni irrisolte che ne impedisco la trasformazione da pratica di nicchia a strumento diffuso, capace di influenzare la nostra comprensione di fenomeni complessi (Vickers e Barrass 2011; Barrass 2012; Nees 2019). La mancanza di strumenti di progettazione condivisi, di protocolli sperimentali standardizzati e di processi di design ampiamente validati sono tra gli ostacoli più citati nella letteratura recente. Può il design, una disciplina che affronta quotidianamente problemi complessi attraverso strumenti specifici, fornire al campo della sonificazione dei dati la conoscenza necessaria per imporsi come uno strumento valido nella costruzione di una più efficace, efficiente e gratificante – migliore - relazione con i dati?

In questo lavoro, ho cercato di rispondere a questa prima ipotesi di lavoro in due modi. Nella parte teorica, ho approfondito la relazione tra la sonificazione e le altre pratiche di design del suono, con particolare interesse ai processi di design, sviluppo e validazione dei risultati che queste pratiche hanno acquisito. Successivamente, attraverso un'analisi condotta su casi recenti, ho esplorato il ruolo dell'intenzionalità dell'autore quale prerequisito di un approccio 'design-driven' alla sonificazione di dati. Al fine di contestualizzare l'intenzionalità – quell'insieme di scelte che il designer fa per rispondere a necessità specifiche, con un obiettivo specifico e in un contesto dato – mi sono rivolta ad esperti di sonificazione al fine di approfondire il legame tra caso d'uso, progettazione, sviluppo ed eventuale validazione dei risultati nei loro progetti. Dall'analisi di queste conversazioni emerge una mappatura dello spazio progettuale della sonificazione, a sua volta utilizzata per definire una serie di blocchi decisionali che formano la base di uno strumento – il data sonification canvas – che ha l'obiettivo di supportare i designer della comunicazione nell'uso della sonificazione dei dati. Il canvas, plasmato e validato attraverso due serie di workshop, chiude la parte teorica e rappresenta uno dei primi risultati tangibili di questo lavoro.

Due azioni progettuali ('Design Actions'), centrate sull'uso della sonificazione per il monitoraggio e la prevenzione degli attacchi informatici alle reti digitali e fisiche, costituiscono l'ossatura della parte pratica. Per le due azioni, sviluppate in collaborazione con la Singapore University of Technology and Design e la società spagnola Ibermática, sono stati prodotti una serie di prototipi di sonificazione, successivamente validati sperimentalmente in ambiente reale grazie alla partecipazione di utenti esperti. La definizione di un processo di progettazione e di un protocollo sperimentale dedicati alla sonificazione di dati, emersi dalle due Design Actions, rappresentano, assieme al design canvas, i due principali frutti di questo lavoro. Infine, questo lavoro apporta nuovi risultati sperimentali alla letteratura sull'uso della sonificazione come strumento di monitoraggio delle anomalie nei sistemi digitali e fisici suscettibili di attacchi informatici. In particolare, il prototipo progettato per la seconda Design Action ha dato vita ad un'applicazione digitale dalle potenzialità commerciali, che è ora in corso di sviluppo.

Questo lavoro vuole mettere in luce, attraverso un approccio design-driven alla sonificazione dei dati, le potenzialità del suono come modalità sensoriale – alternativa o complementare alla più diffusa modalità visiva – capace di contribuire in maniera efficace al processo di comunicazione e comprensione del mondo sempre più complesso con cui, sempre più spesso, ci relazioniamo attraverso i dati.

Index

Introduction

This dissertation was born out of skepticism. In 2017 I was in the process of organizing my return to Europe after five years in the South-East Asian island-state of Singapore. I had moved to Singapore in 2012, bringing with me the sound branding agency I founded in Italy back in 2008. With a master's degree in philosophy and a background in Western classical music, I was consulting for international brands on how to integrate sound into their communication strategy. My team was designing sounds for interaction with physical and digital products in public and private spaces and communication campaigns. A lot of my experience in the design of sound came from experimental art. Since the start of the new millennium I had been active in the Italian sound art scene, experimenting with the first interactive installations that used sensors to engage the public in the creative process and the art experience. I also co-founded the English-language blog www.sounDesign.info, which over time became a privileged viewpoint on the world of sound communication and had run several initiatives to raise awareness on the importance of the soundscape - the sonic environment we are immersed in during our daily life - with children, adults and visually impaired people.

In 2015 I was offered the post of director of marketing in a tech start-up which was developing audio tools for real-time synthesis and processing of sound on web browsers. There is where my first-hand relationship with data began. Through my daily work routine, I became familiar with a continuous incoming flow of data as our go-to-market campaigns rolled out. This included data from Facebook, Google Ads, Instagram, as well as data on engagement, clicks, unique visitors, data from market research and data collected by our apps… Inevitably I started thinking about data sonification - the representation of data through sound - and whether this could be a viable way to support analysis, gather insights and, ultimately, increase knowledge about the phenomenon behind the dataset. I knew very little about the practice of data sonification. I had been asked by a friend to review the thesis of one of his master's students who was trying to represent real-time data on the evolution of user sentiment on Twitter using sound. Behind the scenes, my friend and I both commented that it didn't really seem to be working: it sounded nice but getting any meaningful information from the melody that was produced was a different matter altogether.

This thesis was born out of disbelief that sound could adequately represent and communicate datasets and, more importantly, generate insights on the phenomenon behind the data. I thought that sonification, as I knew it, was not working. What I did not know – and thought was worth investigating – was why. Was sonification doomed to failure? Or was its inadequacy due to specific circumstances that, once identified, could be solved?

Three years down the road, our society has become increasingly and exponentially data intense. We are witnessing an expansion in the tools used to represent this unprecedented amount of data, for which traditional visualization techniques are no longer enough. The practice of data sonification – or simply, sonification – is gaining momentum, as more scientists, journalists, artists and designers integrate it into their toolsets. This thesis presents the results of an investigation into the role of the designer in the transition of sonification from a niche practice for scientific analysis into a medium of mass communication. In doing so, I present the results of two Design Actions where sound is used to represent data for the real-time and pseudo real-time monitoring of anomalous behavior in digital and digital/physical systems and I propose a data sonification canvas as a specific tool for designers who wish to integrate sound into their data representation tools.

Setting the compass. Definitions of data sonification

In her entertaining and compelling 'Lobbying for the Ear: the public fascination with and academic legitimacy of the sonification of scientific data', Alexandra Supper dedicates a full chapter (2012 pp. 75-116) to a historical review and exploration of the various successive definitions of data sonification since the foundation of ICAD (International Community on Auditory Display), in 1992. In the Proceedings of the first ICAD meeting, published in 1994 (Kramer 1994), a first attempt to define and delimit the boundaries of the new discipline is given by Scaletti (1994). Scaletti defines data sonification as "A mapping of numerically represented relations in some domain under study to relations in an acoustic domain for the purposes of interpreting, understanding, or communicating relations in the domain under study." In 1991, she had already provided a working definition of data sonification in her pioneering work 'Using sound to extract meaning from complex data' where sonification is described as addressing the "general problem of how sound can be used to assist the human analyst in the interpretation of a wide variety of data" (Scaletti and Craig 1991). In a recent contribution Scaletti describes data sonification as "a mapping from data generated by a model, captured in an experiment, or otherwise gathered through observation to one or more parameters of an audio signal or sound synthesis model for the purpose of better understanding, communicating or reasoning about the original model, experiment or system." (Scaletti 2018). These definitions focus, on the one hand, on the acoustic nature of sound: in sonification, different parameters of the audio signal (pitch, amplitude, timbre and so on) can be mapped i.e., related to, different parameters of the 'model' i.e., the real-world phenomenon to which sonification refers. On the other hand, all the definitions take into account the purpose of sonification which is, in the 1994 definition as well as in the most recent 2018 definition, that of interpreting data, understanding a phenomenon and/or communicating it to an audience. A third element, which tends to disappear in later definitions of sonification, seems to limit the field to data that can be represented numerically, thus excluding qualitative data or data that is represented semantically (such as data retrieved from the web, that is currently a large area of investigation in the representation of data).

The two poles of 'sound as a parametrizable acoustic phenomenon' and the 'purposes of interpretation, understanding and communication' are maintained in one of the most (if not the most) quoted definitions of data sonification until today. The definition is given by Kramer et al. in the seminal 'Sonification Report' (1999), a collaborative effort compiled for the U.S. National Research Council by the members of the then recently founded ICAD with the goal of... Here, sonification is defined as "the use of non-speech audio to convey information. More specifically, sonification is the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation". When compared to Scaletti's definition, we have two new significant elements. On the one hand, the mapping between relationship in the data and relationship in the acoustic phenomenon seems to lose objectivity with the introduction of the subjectivity of human perception. As it is defined here, the relationship between data (objective) and the perceived characteristics of the acoustic signal (subjective) opens the door to what has been referred to as the mapping problem in sonification i.e., the conflict between a (supposedly) objective reality (the data) and its subjective interpretation by a human listener through sound. On the other hand, the authors of the report took the decision to limit the raw material available to the sonification designer by explicitly excluding speech[1].

I will return to these topics later. This definition, which has survived until today, is not exempt from criticism. Efforts have been made to change it in order to overcome both its limitations (for example, the exclusion of speech audio) and its problematic consequences (for instance, the emphasis placed on the subjectivity of human perception). Hermann, for instance, highlights the contribution of speech and in particular the prosodic attributes of speech as a “valuable element in auditory display” (Hermann 2002, p.23). In 2008, he proposes the introduction of several criteria to assess what can be rightly called a sonification (Hermann 2008), further delimiting the field of application to acoustic representations that “reflect objective properties of data”, in which the transformation from data to sound is systematic, results are reproducible and a same sonification system can be used with different datasets i.e., it is generalizable. On an opposite note, Barras had already proposed a definition of sonification in 1997 as “the design of sounds to support an information processing activity” (p.30) thus emphasizing the importance of sound design in the sonification process while keeping the boundaries open to any activity which involved the production of information. Vickers and Barrass reiterate that “the design approach can allow sonification to become a mass medium for the popular understanding and enjoyment of information in a non-verbal sonic form.” (2011, p.145). While this definition explicitly positions sonification within the realm of design practices, it still retains the limitation on the usage of verbal (speech) sounds as raw material for the representation of information. Last but not least, Worrall defines sonification as “the acoustic representation of data for relational interpretation by listeners, for the purpose of increasing their knowledge of the source from which the data was acquired.” (2009, p.2-4) thus highlighting the difference between the dataset and the phenomenon which data represent and

1. I was unable to find an explicit justification for the exclusion of speech in data sonification. Even if the reasons can be guessed (for example, the human voice has a strong psychological connotation that makes it a very special acoustic material to manage) it would still be interesting to have an explicit formulation for this exclusion.

which is the ultimate goal of sonification as a meaning-making process for the listener. Moreover, once again, it strikes the interpretative (thus subjective) nature of the acoustic representation.

In 2015, in 'Embodied Sonification' (p.5) Roddy attempts a definition from an embodied knowledge perspective. Data sonification is, according to Roddy, "the systematic data driven generation of non-speech sound in order to communicate information about a data source to an embodied listener, who is tasked with perceiving the appropriate meaning(s) within, and/or assigning the appropriate meaning(s) to, that sound." As Figure 1 illustrates, an historical perspective on the definition of data sonification highlights an oscillation between two poles: one that emphasizes the role of the listener, her goals and her role in the definition of the relationship between the phenomenon and the sounds that represent it; and one that sees data as objective entities the properties of which can be mapped to measurable parameters of sound as a physical entity.

Interestingly, in some of the definitions (Scaletti, Kramer), the viewpoint and goals of the listener - interpreting, understanding - seem to implicitly coexist with the goals of the designer – facilitating communication – whereas in others (Worrall, Barrass) only the point of view of the listener or designer (Roddy) is taken into account. This oscillation is, perhaps, the epiphenomenon of a deeper divide within the sonification community between a perspective that sees sonification as an independent means of communication with a general audience and one that sees sonification as a tool in the hands of expert users for scientific analysis.

Starting the journey. The data sonification landscape

The outer landscape - To provide the reader with a first orientation, I used Seealsology[2] (Ricci et al. 2015) to gauge how different fields of sound design are connected to sonification In particular, Figure 2 shows how the Wikipedia page 'sonification'[3] connects to 'computer music', 'sound art', 'sound design', 'sonic interaction design' and 'soundscape' studies.

Only the pages for the macro areas of sound art and computer music have a direct connection to sonification, which is in turn character-

2.
Seealsology is a tool developed by Density Design Lab, Politenico di Milano and the Médialab Science Po that allows to quickly explore the semantic areas connected to any Wikipedia page.

3.
I chose the page 'sonification' instead of 'data sonification' as the latter, when associated to the same pages, gave no results.

ized by the pages of 'auditory display', 'music and artificial intelligence' and 'non-speech audio input'. Sonic interaction design – "the study and exploitation of sound as one of the principal channels conveying information, meaning, and aesthetic/emotional qualities in interactive contexts" (Wikipedia 2021) - which might be thought of as a complementary field to data sonification, shows no direct connections. Neither does 'sound design', a vast field of heterogeneous practices which is not quite reflected in the Wikipedia page. I will investigate the relationship between data sonification and sound design as part of this thesis. It is interesting to note that in the Seealsology network shown in Fig.2 they are not connected at all. On Wikipedia sonification positions itself as a sort of middle ground between (sound) art and (music) computing, perhaps mirroring the oscillation between hard science and humanism which we saw in the evolution of the definitions of data sonification (see Fig.1).

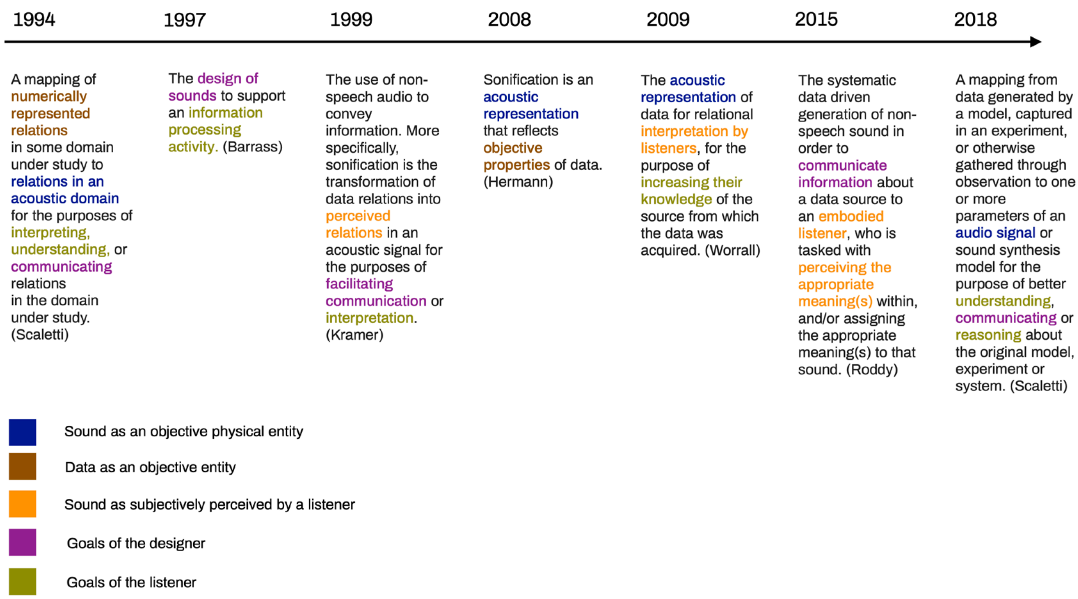


Fig. 1: definitions of data sonification 1994-2018.

The inner landscape - The practice of data sonification is generally categorized according to three techniques used to translate data into sound: audification, parameter mapping and model sonification. Each category is extensively described in 'The Sonification Handbook' (Hermann, Hunt and Neuhoff 2011). For the purpose of this introduction, I will briefly recap their definition:

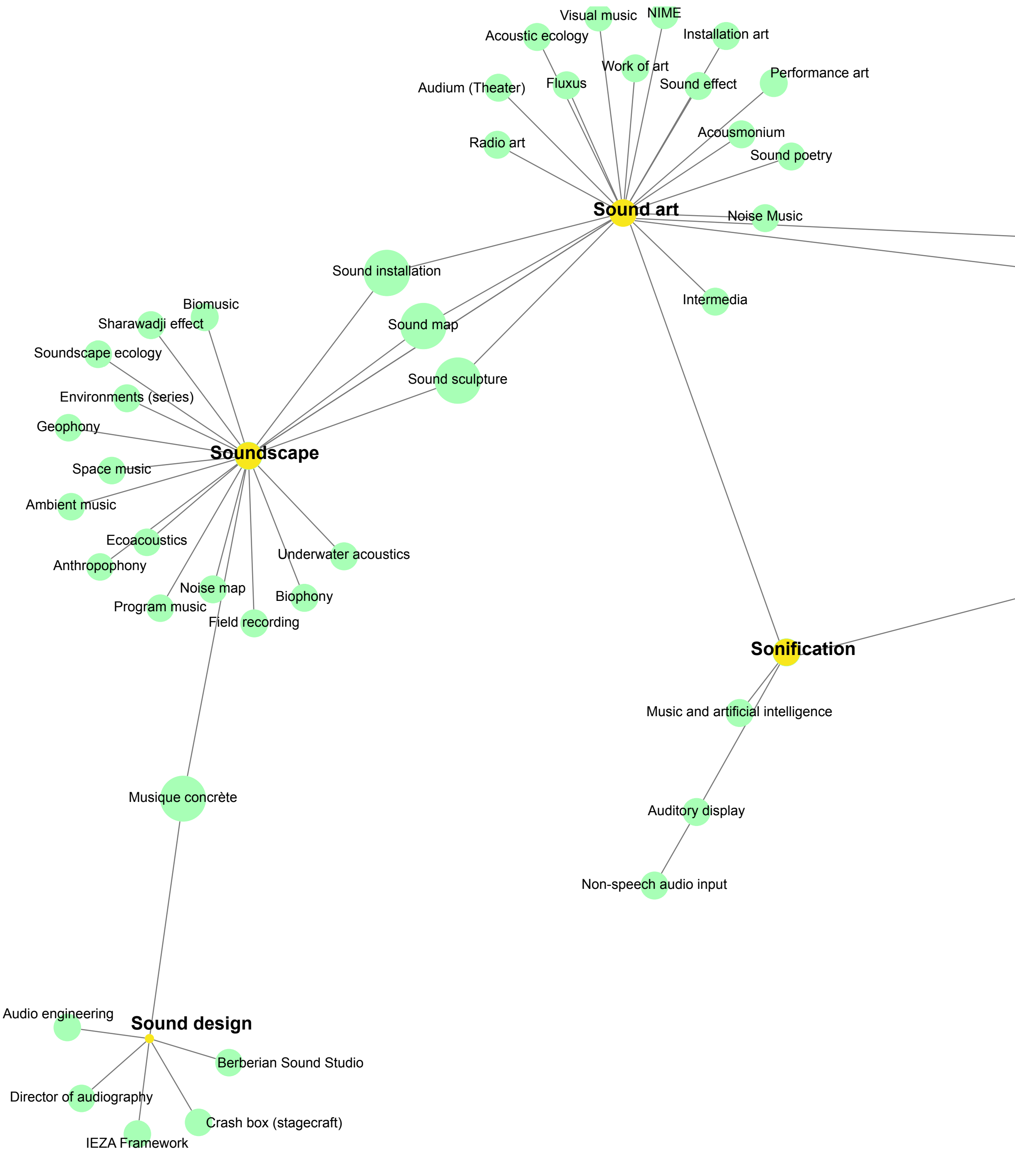
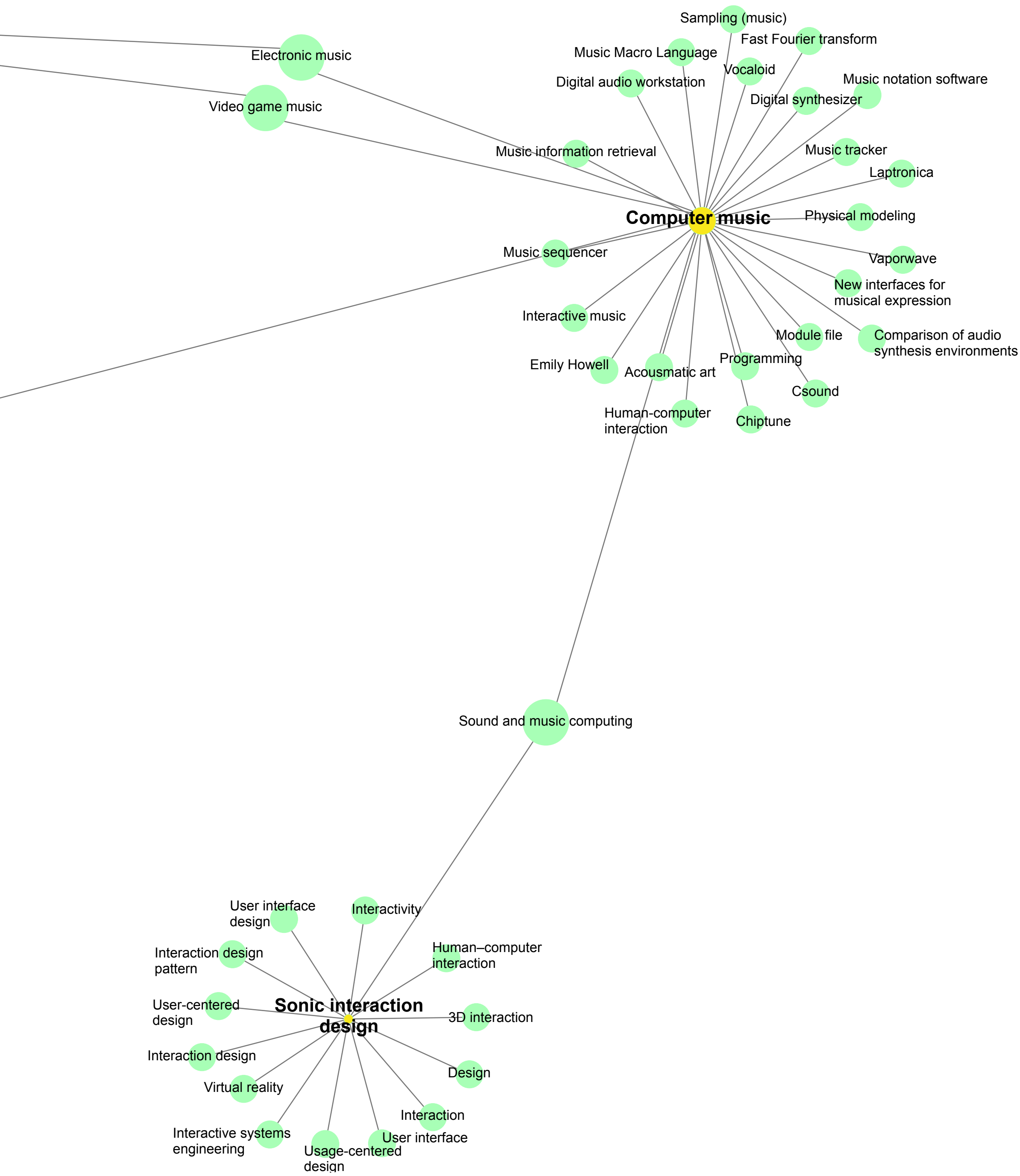


Fig. 2: ‘sonification’ Seealsology network.


Sampling (music)
Fast Fourier transform
Music Macro Language
Vocaloid
Digital audio workstation
Music notation software
Digital synthesizer
Electronic music
Video game music
Music information retrieval
Music tracker
Laptronica
Computer music
Physical modeling
Music sequencer
Vaporwave
New interfaces for musical expression
Interactive music
Module file
Comparison of audio synthesis environments
Emily Howell
Acousmatic art
Programming
Csound
Human-computer interaction
Chiptune
Sound and music computing
User interface design
Interactivity
Interaction design pattern
Human–computer interaction
User-centered design
Sonic interaction design
3D interaction
Interaction design
Virtual reality
Design
Interaction
Interactive systems engineering
User interface
Usage-centered design

- Audification. Perhaps the most rudimentary sonification technique, audification consists in “The direct translation of a data waveform into sound” and as such, it requires the listener to “simply listen to the event or listen to data generated by the event.” (Walker and Kramer 2004, p.152). The translation process would often require “that the data wave be frequency shifted into the audible range for humans or time shifted (slowed down or sped up) to allow appropriate inspection by the listener” (ibid).
- Parameter Mapping. In what seems to be the most frequently used technique (Supper op. cit., p.15), data dimensions are systematically mapped to specific sound parameters such as frequency, amplitude, timbre and rhythm in a procedure that echoes the definition of data sonification given by Scaletti.
- Model-Based Sonification. This technique introduced by Hermann at the end of the 1990s (Hermann and Ritter 1999) conceptualizes sonification as a dynamic framework that translates data into sound in response to interaction from a user. In Hermann’s words, a model sonification “does not make sound at all without external interactions (in other words: the data is used to build an instrument or sound-capable object, while the playing is left to the user).” (Hermann 2008 cit., p.1).

Research Questions

As I confessed earlier, I started my doctoral research with the goal of challenging as objectively as possible my own skepticism around the possibility of using sound to represent data in a way that ‘really worked’ for the final user. The first year of my doctoral research was dedicated to a preliminary research which included the review of the official definitions of data sonification and the positioning of sonification in the landscape of other sound-related fields. Along with my personal experience as a sound designer and my training in electroacoustic music and sound design for film, this preliminary research helped me identify challenges and opportunities around three main areas of research interests that I expressed in three research questions. Firstly, I understood that, in order to answer my overarching question on the potentialities of data sonification as a real-world application, I had to reflect on the contribution that I, as a self-defined sound designer, would bring to the field when tasked with using sound to represent real-world phenomena through data. In doing so, I would have to map the specificities that a design-driven approach to sound entails and define it in relation to other fields that are tangential to data sonification such as, for instance, computer music, soundscape studies, Sonic Interaction Design (SID) and so on. The formulation of the first question that emerged from this process was simple and direct:

> <u>What is the role of the (sound) designer in data sonification?</u>

Following a more in-depth investigation, this very broad first question revealed the implicit assumption, that is largely based on my personal story, professional background, aesthetic preferences and general approach to the world of sound, that a design-driven approach to the sonification of data would produce 'better' sonifications. I further described 'better sonifications' as sonifications that would be more engaging for the final user - thus building a more compelling case for the use of sonification in the real-world - and more efficient i.e., able to support users in reaching their goals. To make the assumption explicit, I added a second line of inquiry to the first Research Question (RQ1):

> RQ1: What is the role of the (sound) designer in data sonification? Can a designerly approach to sonification make the difference in creating better (more efficient and engaging) representations of data?

As the reader might expect, the reference to data sonification improved by (sound) design opened the door to a second research question which posited the need to explore and define a specific evaluation protocol for assessing this improvement. This need was also grounded in literature since for three decades scholars have repeatedly warned about the lack of experimental and real-world evaluations of sonifications. Through the second RQ I planned to explore, validate and define a possible experimental protocol for real sonifications - in particular those specifically designed during my doctoral research - but sufficiently generalizable to be applied to other projects. Additionally, through field research such as interviews with experts and workshops, the exploration and identification of the specificities of sound design for data sonification led to the definition and conceptualization of a design tool, the sonification canvas, which helped answer the second part of RQ2 and forms the backbone of the theoretical part of this dissertation.

> RQ2: How can we evaluate if they are better? Can we frame a design methodology to approach data sonification projects from prototyping to testing?

Over the course of the first year, I was presented with the opportunity to apply data sonification to the monitoring of cyber-attacks on a digital-physical system (a water distribution network) and on a fully digital system (an Internet network), in collaboration with two international partners. I called these two cases Design Actions (DA). During the second year of my doctoral research I conceptualized, designed and implemented two prototypes (one for each DA) for monitoring and detecting anomalies due to cyber-attacks on digital-physical networks. The two DA were tested by domain experts in a real- or semi-real-world (partly due to the consequences of the COVID-19 pandemic, as I will illustrate in Chapter 4) context in what I considered an opportunity to answer the second RQ. They also formed the knowledge basis for exploring the third area of investigation (RQ3):

> RQ3: Can data sonification represent the complexity of digital and digital-physical systems in order to help expert users detect and prevent anomalous behavior? Can it have an impact on their daily activity?

The second part of RQ3 would later bring me back to the RQ2 since the potential impact of sonification on the activity of expert users has to be evaluated in order to answer the question. The structure of the research is therefore somewhat circular as the theoretical investigation intertwines with the practical development of the Design Actions in a process that is reminiscent of programmatic research (Bang and Eriksen 2014), as shown in Figure 3.

In so-called programmatic research, an overarching research challenge is defined and framed in terms of specific research questions while systematic actions (experiments or, in the case of this work, Design Actions) are carried out throughout the entire investigation to iteratively guide the program from the status quo to the new framework (Binder and Redström 2006) Broadly speaking, programmatic research includes both theoretical research - in order to answer the research questions and transition to a proposed new framework for the field of study - and practical research punctuated by specific actions and experiments. In this thesis, I presented a specific design process and an original design tool, the sonification canvas, as the main results of the theoretical investigation, while results of the application of sonification to the monitoring and detection of anomalies in digital and digital-physical systems in the context of cyber-security are presented as outcomes of specific actions and experiments.

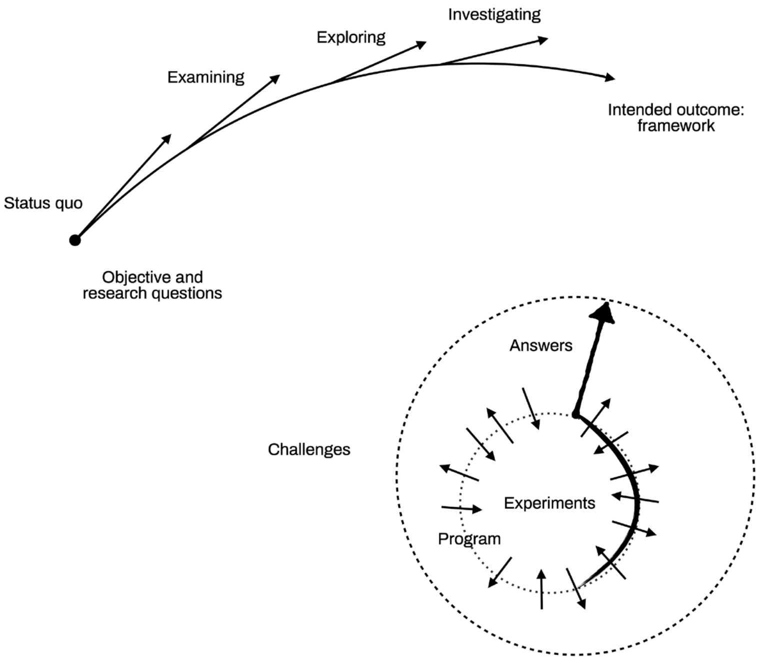


Fig. 3: scheme of programmatic research (adapted from Bider and Redström 2006).

Methodology

Four methodologies were used to address the three RQs over the course of the doctoral research. In particular:

M1: Literature Review. Both from primary and secondary sources. Primary sources include the Proceedings of the International Conference on Auditory Display (ICAD) which celebrated its 25th year of activity in 2019.

M2: Qualitative research. In the form of a semi-structured interview with domain experts, qualitative research was applied as an experimental method in the evaluation of both design actions and as a research method in the theoretical part. Through two workshops, qualitative research was also used in the evaluation of the sonification canvas, which will be described in Chapter 2 of this thesis. Finally, to investigate intentionality in sonification I used case studies analysis as a qualitative research method.

M3: Quantitative research. Quantitative testing, in the form of a task-based performance assessment, was part of the experimental protocol applied to the evaluation of both design actions. Results are presented in Chapter 3 and 4.

M4: Design Actions. Two design actions were conceptualized, designed and prototyped in the course of the doctoral research. The prototypes, which involved programming and coding as well as sound design, were developed with the help of external collaborators. In particular, Ginevra Terenghi, a communication design master's student at the time of the investigation developed the software part of the first DA prototypes. Prof. Damiano Meacci(Conservatorio di Firenze 'Luigi Cherubini') developed the software infrastructure and the final application for the second DA.

Research Strategy and Thesis Structure

Before describing the structure of this dissertation, I want to share with the reader a timeline of my doctoral research from November 2017 to early 2021. In Figure 4, light blue blocks represent mainly theoretical investigation while red blocks represent practical research. The first part of my doctoral research was dedicated to the formulation of the three research questions mainly through literature review, course attendance and lab activity. DA1 was developed during the end of the first year and the first half of the second year. The activity of conceptualizing, prototyping and testing both Design Actions involved both theoretical and practical investigation, as Figure 4 illustrates.

During the second year, and while working on DA1, I carried out a 12- months internship with the Artificial Intelligence Innovation Team of i3B, Ibermática, in Spain. The second Design Action (DA2) is the result of this internship. As DA2 took place in a real industrial environment within a commercial company, I had less control over the timeline for the design and development process. As highlighted in Fig. 4, R&D time for DA2 took longer than DA1. In the midst of the prototyping phase, which should have been followed by an experimental phase, an unforeseeable event burst into our lives as the entire world went into lockdown due to the COVID-19 pandemic, which is still ongoing at the time of writing. The prototyping phase came to a halt as did plans for testing in a real working environment. R&D continued from home, while I focused on developing the theoretical part of my research and notably, the definition of a design framework for sonification which later became the sonification canvas. Towards the end of 2020, despite the consequences of a second wave of COVID-19, I was able to resume work for the DA2 and apply the prototype to a real-world testbed. Due to restrictions on access to industrial laboratories, the experimental phase had to be severely reduced compared to the original plan and postponed until the very last weeks of my PhD project.

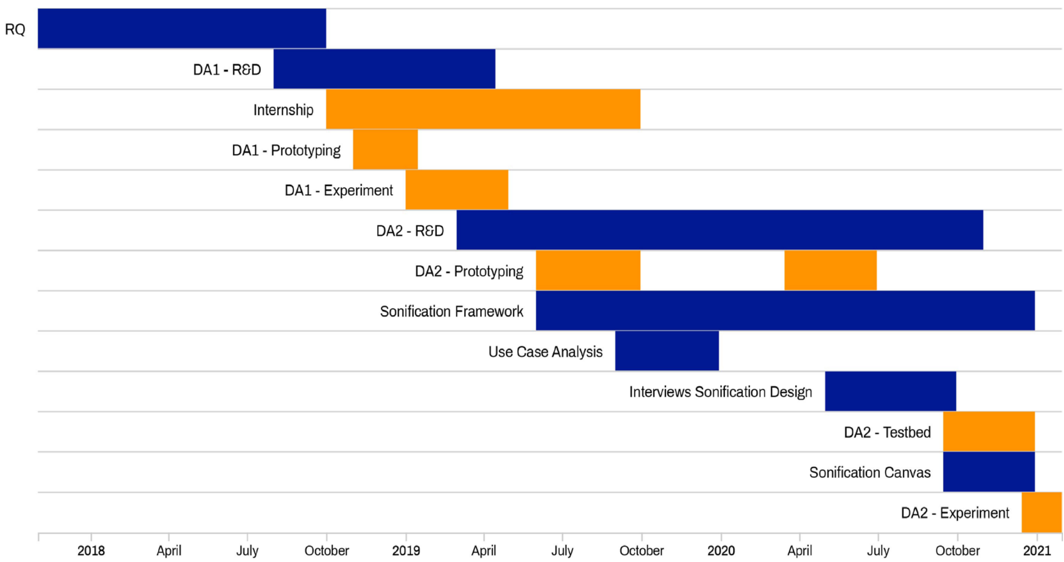


Fig. 4: GANTT chart of the doctoral research, 2018-2021.

Figure 4 illustrates the evolution of the research process which is consistent with the programmatic research approach illustrated above. An overarching continuous theoretical investigation focused on the development of a new framework for the design of sonifications is punctuated by applied design actions which are used to stimulate as well as to validate theoretical reasoning.

Another point worth noting is that I only focused on the development of the data sonification design tools and processes at a late stage in the doctoral research, while research applied to real-world cases of sonification design through the design actions was carried out mostly in the first half of the project. As a consequence, and perhaps ironically, the design of the two Actions does not follow the approach which I will propose as one of the main results of this thesis. This partial inconsistency is due to many factors, including opportunity (the timing of design actions was dictated by the availability of partners) and perhaps even the natural evolution of the research process as the maturity to face theoretical questions comes after one has contended and struggled with practical, hands-on work. Finally, it is also partially due to the internationally unprecedented situation of the coronavirus pandemic and the radical change this provoked in the way we all work.

Nonetheless, in the presentation of the results, I decided to follow what I believe is a more coherent and consistent logical order, with the theoretical part preceding the presentation of the practical investigation and the conclusive part in which the first and second section merge and theory is grounded in practice. The first part is structured as follows: from an introduction to the field of sonification, in which I delimit its boundaries, identify current issues and delineate a possible role for design, I move towards the presentation of how the sonification canvas was developed through qualitative research which includes a case – studies analysis, interviews with experts and two workshops. The practical part is fully dedicated to ground the theoretical research in the practical work carried out in the field of anomaly detection to monitor and detect cyber-attacks on digital and digital-physical systems. This second part describes in detail the development, design and validation of the two DAs. The conclusive section – Part III – presents and discusses the results obtained in three areas: the proposal of a design-driven framework for sonification, which includes the sonification design canvas and the definition of a work process grounded in the DAs; the proposal of an experimental framework for the validation of sonifications in a real-world environment, also grounded in the DAs; the discussion of the experimental results obtained through the DAs in the field of anomaly detection in digital and digital – physical systems for cyber-security. While working on the methodologies illustrated in this Introduction , it struck me that its three-part form in which two main themes (the theoretical and the practical research) are presented greatly resembled another classical structure which I cherish from my years spent at a Music Conservatoire: the forma sonata (the Sonata Form) - perhaps the most classical structure of Western culture. As an homage to my foundational years in the world of classical music, the titles of each part and the titles of the chapters mimic the names of the three-sections and two-themes classic sonata form. Following the tradition of the 18th century classical music, these titles will be kept in Italian: the esposizione (exposition, Part I) will present the primo tema (first theme, the sonification canvas), while the sviluppo (development, Part II) will see the presentation of the secondo tema (second theme, sonification for anomaly detection) in two variazioni (variations, DA1 and DA2). The ripresa (recapitulation) will repropose the two themes now in their maturity, as we move towards the conclusions.

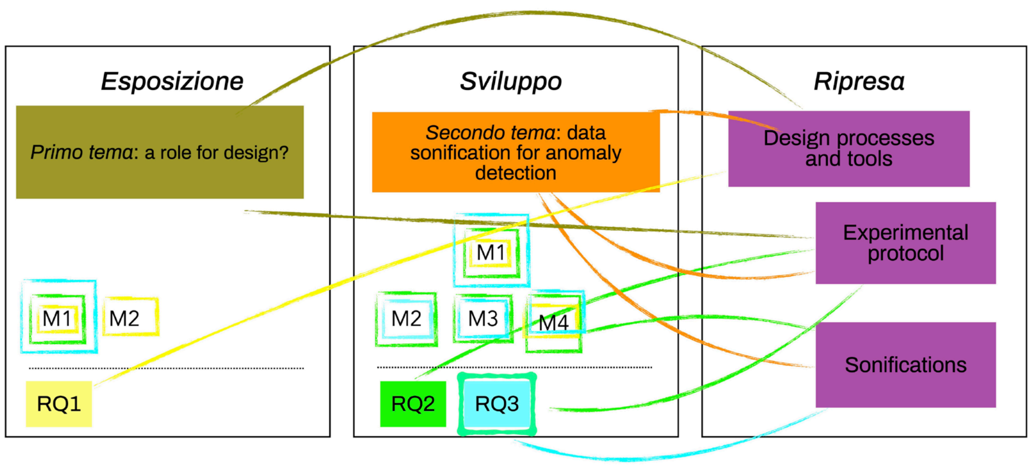


Fig. 5: Outline of the thesis.

The diagram in Figure 5 summarizes the structure of the thesis and shows how it connects to the three research questions and the four methodologies illustrated in this Introduction.

A doctoral research project is obviously not as linear as a sonata. To account for all the (mostly fruitful) deviations from the main path, Fig. 5 includes a second layer, where themes, research questions and methods intertwine to lead us to the main results presented in the thesis: tools, processes and experimental protocols for the design of sonification projects, experimental findings and sonification prototypes for the two cases developed through the design actions.

Glossary

To facilitate the reader not familiar with acoustic parameters of sound, I include a glossary of the terms most frequently used in this research. All the definitions are taken from (Everest, 1996), translated by the author of this work.

Pitch: the subjective perception of the frequency of a sound.

Frequency: it is the value of the speed at which a sound signal is repeated in a second. It determines how we perceive the pitch of a sound: the higher the frequency, the higher the pitch.

Timbre: the quality of a sound, determined by its harmonic components. Different musical instruments, or voices, are perceived having a different timbre even when their pitch is perceived as identical.

Amplitude: it is the value of the amplitude of oscillation of a sound wave. It is perceived as the volume of a sound: the higher the amplitude, the louder we perceive the sound.

Loudness: the subjective perception of the amplitude of a sound.

Hertz: it is the unit used to measure the frequency of a sound signal. The symbol is 'Hz' and it corresponds to the number of cycles that a sound wave does in a second.

Decibel: it is the unit used to measure the intensity of a sound signal. The symbol is 'dB' and it corresponds to one tenth of a bel (B).

Part I.
Esposizione.
Defining a design space for data sonification.

Part I.
DEFINING A DESIGN SPACE FOR DATA SONIFICATION

The first chapter – Orientation – aims to set the register of the first part, as a composer would choose the register i.e., the range of her compositional space before venturing into the details of the work. It opens with a review of the state of the art for the practice of data sonification. In the first section, ‘Sonification, a discipline in flux’, I present a series of ‘divides’ which compare the status of sound-based practices to visual-based practice, thus highlighting the main challenges data sonification is facing today. The last three points – what I call the ‘design thinking divide’, the ‘mapping problem’ and the lack of experimental validation of sonification artifacts - open the door to a deeper investigation of the two first research questions in the frame of design in the following section. In ‘A role for design’, I expand on the assumptions behind each research question. First, I explore and define contemporary sound design practices such as sound design for film, sound branding and Sonic Interaction Design.

Secondly, I present existing structured work processes in different sound design fields and assess their similarities with design practices. The third part of the section is dedicated to the presentation of existing experimental protocols for the validation of data sonification projects and the preliminary evaluation of protocols borrowed from design research.

In Chapter 2 - Exploration, I choose the key (should it be minor, major, or a combination of the two?) of the theoretical research which will lead to the definition of the first theme – the data sonification canvas. The chapter explores the criteria that define a data-driven approach to sonification. The first section explores the criterion of intentionality in data through the evaluation of five recent cases. Intentionality is a necessary but insufficient condition for assessing a designerly approach to sonification. The following section – 'Ritorno alle cose: the daily practice of sonification' - goes 'back to things' in a neo-phenomenological approach that aims to evaluate the daily practice of sonification design through interviews with authors of recent sonification projects. In this section, I outline and define a series of decision- from the qualitative analysis. These ‘decisions’ inform a first draft of the sonification design space which will become, in the following chapter, the sonification canvas.

In the first section of Chapter 3, ‘Building the Blocks’, I go through the definition of the canvas’ elements. The sonification canvas prototype thus obtained is validated through a workshop in which communication designers use the canvas to categorize existing data sonification cases. Results of the workshop are used to inform a second canvas prototype which is also validated through a workshop in which designers use the canvas as a tool for the conceptualization of real sonification projects. The process and the results of both workshops are described in detail and the final form of the sonification canvas is presented to the reader.

CHAPTER 1 —
Orientation: setting the register

Soundtrack: *J. S. Bach, The Goldberg Variations. Glenn Gould.*

Sonification, a practice in flux

The practice of using sound as a structured method to understand, represent and communicate insights into data sets can be traced back to the 1990s. ICAD, the International Community on Auditory Display, was founded more than 25 years ago as a forum for sharing practices, discussing and defining sonification as a discipline. The auditory representation of data was, at that time and for a large part of the past two decades, mainly focused on the representation of scientific data for the purposes of expert analysis (Kramer et al. cit.) and as an alternative or a complement to the visual representation (Scaletti and Craig cit.; Hermann, Hunt & Neuhoff cit.).

Over the years, the use of sound to represent data has covered various fields. In seismology, Dumbois (2001) used audification to represent and analyze earthquakes, while Ballora and colleagues (2000) applied the same technique to the sonification of heart rates. Nesbitt and Barrass (2002) combined visualization and sonification for the analysis of stock market data. Dayé and de Campo (2006) used parameter mapping sonification to represent sequential data (temporal and spatial) in social sciences. In some cases, like in the work of the astronomer Wanda Diaz Merced (2013), sonification has been used to support visually impaired scientists in their investigation, as an alternative to data visualization.

On some occasions, digital tools such as the Sonification Sandbox by the Sonification Lab at the Georgia's Institute of Technology (Walker and Cothran 2003) and SoniPy (Worrall et al. 2007) were developed and shared with the community with the goal of providing researchers with general-purpose instruments for applying sonification to their datasets, so they would not have to develop new ones. In the case of the SonEnvir project (de Campo, Frauenberger and Höldrich 2004), sonification frameworks were developed to support social scientists in the use of sonification and were tested through workshops. Moreover, basic research was carried out with the aim of identifying generalizable mapping strategies both by experts in sonification (Walker and Kramer 2005; Roddy cit.; Vogt and Höldrich 2010; Rönnberg 2017) and by researchers not directly involved in the field of auditory display (Ballatore, Gordon and Boone 2018).

In recent years, sonification seems to have expanded its reach beyond the original scientific community. With the exponential increase in both the volume and impact of data in our society, the need to complement an overloaded visual channel with other sensory modalities has increased and the use of sound to represent data has gained momentum. A wide variety of data-driven use – ranging from artistic expression to advocacy and data journalism - has been labeled (both in academic literature and mainstream communication channels) as data sonification, alongside and beyond any official definition

The newly launched Data Sonification Archive (Lenzi et al. 2021) has about 150 entries to date, 20% of which were produced in 2020, with 17 of these dedicated to the COVID-19 pandemic. Commercial applications such as Microsoft Power BI (Eldersvelt 2019) include tools for the sonification of dashboard data, while Google News Initiative launched its own browser – based sonification tool, Two Tone (Cairo 2019), in collaboration with the data visualization agency Datavized Technologies[4]. Two Tone, "A free web app to turn data into sound and music", can be used "for understanding data through listening. It makes data – claims twotone.io homepage - more accessible." A dozen sonification projects produced with Two Tone are featured on the Google

4. https://datavized.com/

News Initiative page. Authors of data physicalizations occasionally make use of sound, along with other sensory modalities, for data representation (Hogan and Hornecker 2016). Investigative journalists are using sonification to engage the public during radio broadcasts (Cory 2015) or as a source of data in the context of forensic investigations on human rights issues (Weizman 2017).

The auditory channel is also integrated as an additional dimension in data visualizations (Sobliye and Mortada 2017) by authors aiming to activate their audiences on socially relevant topics. At the same time, there seems to be an increasing debate both within the community and in the larger design community around emerging themes, such as the potentialities of sonification for facilitating decision-taking in cyber-security (Lenzi et al. 2019; Axon et al. 2019; Vickers 2014); the role of aesthetics in sonification, where the emphasis is on the relationship with the listener (Roddy and Furlong 2014; Vickers and Barrass cit.; Barrasss 2012); the need for sonification to strengthen its roots in design theory (Nees 2019) and finally, the framing of sonification as a design-driven practice in which the author intentionally places sound at the center of the communication process (Lenzi and Ciuccarelli 2020).

Still, the real impact of sonification on everyday life in terms of the improvement or transformation of our daily relationship with data is at best extremely limited since the “legitimacy and usefulness” of the practice is “still controversial” (Supper cit., p.10). Indeed, despite the exponential increase in the production and consumption of data at all levels of society, outside the specialized community data sonification is still largely viewed as a means for catching the audience’s attention in popularization efforts rather than a medium that can be used to make sense of data, communicate information and build knowledge (Masud et al. 2010). The reasons for this underrepresentation are certainly manifold. In the following paragraph, I will share with the reader a non-exhaustive list of what I believe are the most prominent current issues of sonification. This list has been compiled based on a review of existing literature, informal conversations with peers and interviews with experts. I use the word ‘divide’ as a metaphor for the gulf sound practitioners sometimes perceive between their field and the visual world based on the uneven distribution of resources and attention (Supper cit.).

While compiling the list, I became aware of a certain inevitability in having to implicitly compare sound with the visual world. The relationship between sound and image is an ancient and complex one. It is believed that the Greek philosopher Pythagoras used to address his disciples (the 'akousmatikoi'[5]) from behind a veil to prevent them from being distracted by the sense of sight and help them focus on his voice - the acoustic medium through which his teachings were shared. It is from this very practice that Pierre Schaeffer, the founding father of electroacoustic music, named acousmatic music i.e., a music composed with "sounds one hears without seeing their originating cause" (Schaffer 1966, p.91-99). A music conceived and produced to be shared with the public from behind the veil of loudspeakers, liberated from the vision of the instrumental gesture of the performer which inevitably ties the sound, in the mind of the listener, to its causal source. In the words of Chion, who first retrieved the notion of acousmatic listening after Schaeffer, "Acousmatic sound draws our attention to sound traits normally hidden from us by the simultaneous sight of the causes" (Chion 1994, p.32). Chion addresses the relationship between sound and the moving image in his seminal work 'The Audiovision – Sound on Screen', which I will have quoted on several occasions in this work. One of the founding mythologies of cinema since the advent of sound is the soundtrack, Chion explains. The soundtrack, as the name says, is implicitly considered an addition to what constitutes the real nature of cinema, the moving image. This is nothing but an illusion, specifically, an audiovisual illusion (ibid., p.5) thanks to which we watch a movie believing we see things that, on the contrary, we only hear. Sound, far from being a mere beautification of the image, changes the way we understand what we see.

5.
Akousmata means "oral saying", so the akousmatikoi were those who listened to the oral teaching of the philosopher. From Wikipedia, the free encyclopedia.

Still, there is a generalized feeling among professionals that the art of sound remains somewhat minor, or is even neglected, when compared to the importance given to the visual medium in our culture. Supper dedicates an entire section of her work (Supper cit., Chapter 3) to exploring the sense of dominance of visual culture and the marginalization of the world of sound the auditory display community is trying to liberate itself from by building an "emancipatory rhetoric that promises to free the ear from its marginalised status within science, to lobby for the acceptance of sound as an authority and let human listening skills unfold their true potential" (ibid., p. 75). The following non-exhaustive list instantiates the most recurring claims and outlines current changes in the divide between the visual and the aural culture.

The cultural divide- As mentioned, the existence of “a cultural bias towards visualization” (Hermann cit.) is a common claim among sound designers and perhaps even more so among members of the auditory display community.

In a 2006 article Dayé and de Campo posit that “Nowadays, Western culture as a whole is a visual culture: It is a culture of seeing, of reading, a culture of scripture and images. [...] Over centuries of philosophical debate, the eye emerged as the only sensory way to the truth, in scientific, theological and social contexts” (cit., p.352). The perspective of visualization becoming synonymous with scientific truth is shared by Latour in ‘Visualization and Cognition’ (1986) as well as in his influential 1979 work with Woolgar (Latour and Woolgar 1979) where the process of attributing predominance to science over other cultural human manifestations is ultimately connected with the laboratory practice of using written representations and written language which become instruments of authority and domination[6].

The technological divide- There are two aspects of what I call the technological divide between visualization and sonification. The first aspect relates to the existence and availability of tools for the production of data sonification projects. The second relates to the availability of devices for the fruition of sonifications. In the first case, as early as 1991, Scaletti (cit., p.3) pointed to the lack of “general purpose real-time sound synthesis hardware” and “An absence of models that would allow sonification systems to run on several different hardware platforms” i.e., the absence of both hardware and software infrastructure equivalent to that existing in the visual world, where graphic design and image processing software as well as mainstream software such as spreadsheets allow even the non-expert user to create quick data visualizations. Additionally, recently developed online infrastructure such as Raw Graphs (Mauri et al. 2017) allows any user (expert or non-expert) to take advantage at no cost of high-quality, customizable data visualization production tools.

Even today, the lack of such ‘off the shelf’ tools often obliges the authors of sonifications to rely on tailor-made software that has to be

6.
This predominance is, according to Latour, mainly based on the power of written language. This makes the voluntary exclusion of speech from the definition of sonification even more mysterious than I stated.

programmed specifically and from scratch every time a certain degree of interactivity between data and sound is required. Notable exceptions are the Sonification Sandbox (Walker and Cothran cit.), a tool that "allows the user to map data to multiple auditory parameters and add context using a graphical interface" motivated by "a need for a simple, multi-platform, multi-purpose toolkit for sonifying data", as explained in the homepage of the project. The recently launched Two Tone (Cairo cit.) is a fully browser-based tool that allows the user to upload data in the form of a text file and create sound and music through the selection of simple parameters. Though Two Tone does not allow for a complex treatment of sound material, it represents a step forward in the dissemination of sonification to a wider, non-expert audience of potential authors.

The second aspect, the fruition of sonification in a real-world context, has been raised by Dayé and de Campo, who highlight that "Standard working environments often turn out to be quite unusable for sonification research, simply because their design follows the cultural hierarchy of the senses" and that "When using speakers, the researcher should be seated in a room alone, since it is likely that room-mates will not accept being subjected to computer generated sounds over an extended period. When, on the other hand, one is using headphones, one makes oneself more or less unavailable to other technical devices that make use of an auditory interface." (cit. 2006, p.360-361). Whereas smartphone technologies are promoting the habit of wearing wireless earplugs at all times, thus opening the door to the daily usage of applications that need sound to communicate, the divide with the amount of hardware and software devices available to the common user for visually-based interfaces and applications in our everyday life remains unbridgeable.

The educational divide- In a recent article, Wirfs-Block et al. state that "sonificators focus on production rather than explanation" (2021 p.3). While this is certainly true and deserves far more attention from the sonification community, it might reflect a wider issue regarding education in the field of sound and music studies as compared to the level of average visual literacy. It is a common claim in the sonification world that while the current school system does, in general, teach pupils to read and produce data visualizations (starting from primary school, where we become familiar with the Cartesian axis) such a foundational education is almost completely lacking when it comes to learning how to decode aural messages, even those used in organized sound structures such as music compositions. It therefore becomes difficult to expect the average listener to be able to decipher messages carried by sound-based artifacts such as sonifications without prior training, even though we take for granted that, with the aid of visual keys, the average reader of any newspaper is able to read and understand data visualizations on a daily basis.

Moreover, there is a divide in the literacy of users of sonifications. We encounter this same divide when it comes to the education and training of future sound design professionals. Indeed, structured education in sound-related disciplines, at both graduate and post-graduate levels is also underrepresented. As a mere example, a quick search on niche.com[7], a popular website for research on schools and the educational offer in North America, presents the visitor with a list of the 38 top master's courses in graphic design while the same search on sound design returns zero options.

7. https://www.niche.com/graduate-schools/search/best-masters-in-graphic-design-programs/

The portability divide- Lastly, it is undeniable that the traditional support for the distribution of human knowledge i.e., written text printed on paper, "hardly begins to meet the requirements of communicating sound." (Dayé and de Campo cit., p.360). The limitations in the distribution of sonification projects (as well as any sound-related artifact) are well known to the sonification community (Supper cit., p.190). Nowadays, digital technologies – starting from the CD first and later personal websites and online archives such as Soundcloud - as well as the digital formats currently used for the distribution of scientific knowledge online, allow for the inclusion of sound works in scientific papers. Nonetheless, a quick research I conducted during the first year of my doctoral project revealed that only four out of 60 scientific articles on sonification had related audio content that was available to the reader.

There are three more aspects of the divide between sonification practices and the world of visualization that deserve a separate focus. They explicitly concern the main recurring themes of this doctoral investigation, that can be traced back to the first two research questions and run as a background thread throughout this work: the need to define shared tools for the design of sonifications as a way to fulfill the potentiality and expand the reach of this practice (what I call the design thinking divide); the issues that arise when we represent data through sounds that are still mainly chosen arbitrarily, “without regards for any ‘natural’ connection to the data represented” (Scaletti cit., p.3), generally referred to as the mapping problem, and the lack of a shared experimental protocol or experimental practices for the validation of sonifications. I will shortly introduce each of them before they are extensively examined in the following paragraphs.

The design thinking divide- In Chapter 4 of ‘Lobbying for the Ear’ (cit.), Supper reports an enlightening episode. During a presentation at the International Conference on Auditory Display (ICAD 2010), to the question on how the data were mapped to sound “The speaker adopts a somewhat defensive tone in his reply: yes, some decisions about how to map and process the sounds had indeed been made, ‘just like you do in visualization’, but the mapping itself was very precise and reproducible.” (Supper cit., p.117). There are many aspects of this statement that would be worth investigating, first and foremost the possible inconsistency between taking specific decisions “just like you do in visualization” and a “very precise and reproducible” outcome of such (I imagine, subjective) decisions. In Supper’s investigation, the tension expressed in this sentence reflects the struggle of the sonification community between the ambition of seeing sonification recognized as a valid scientific method and the need to leverage sound design and sound art practices which could better account for the creative decisions taken during the sonification process. The status, nature and historical evolution of a design science that can both account for subjective decisions as well as reproducible and generalizable results has been a hotly debated topic in the field of design research since at least the 1960s. It is not the goal of this dissertation to enter this debate but I cannot avoid recapping some of its core themes in relation to the current status and nature of the field of data sonification.

Outside select circles of contemporary design research, it is largely taken for granted that design is today a structured discipline which can be taught, learnt and practiced at many different levels of human activity, within academia and industry. Its reach crosses a variety of fields, from the design of consumer products, to services, to communication artifacts - which range from graphical supports, digital interfaces, real-world experiences and even public participation and governance processes. 'Design thinking', a rather over-used expression which was introduced as a commercial methodology in the 1990s and has since been applied, consistently or otherwise, to innumerable fields, does actually try to break down the mainly tacit creative process through which designers approach the production of new knowledge. This methodology, schematized in Figure 6, covers, with minor differences, the process designers of different fields – from architecture to fashion, user experience, service or communication design, to name but a few - engage with when addressing a new creative challenge.

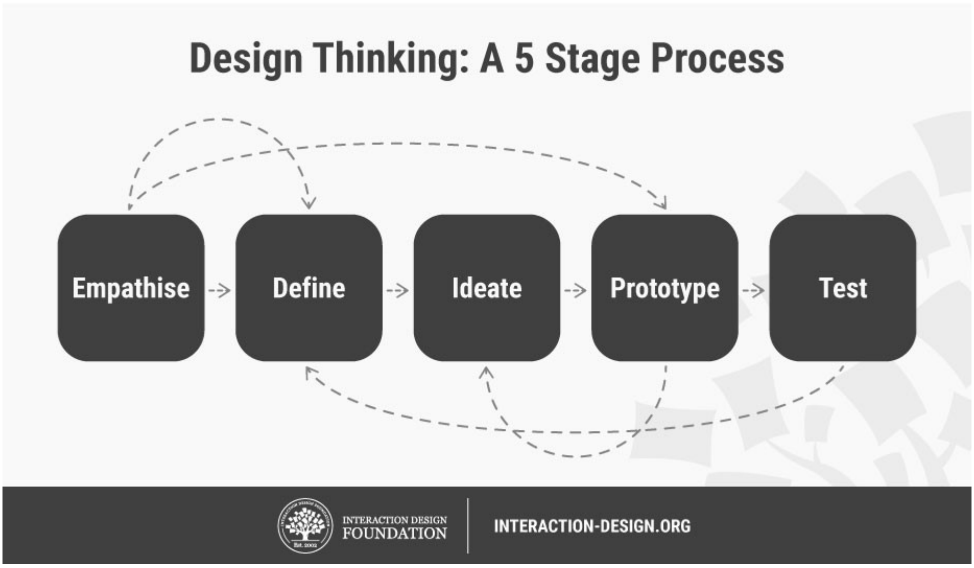


Fig. 6: Design Thinking diagram, Interaction Design Foundation.

I will try, without the ambition of being exhaustive, to tackle the assumptions behind the designerly way of doing and knowing (Cross 1982) in the upcoming section. It is worth noting here that sound is completely absent in the definition of the 'four broad areas' of design presented by Buchanan, in his seminal 'Wicked problems in design thinking' (1992, p.9), while, as Figure 7 illustrates, symbolic and visual communication occupies the first place (followed by material objects, activities and organizational services, and complex systems).

To my knowledge, a structured, codified approach to the design of sound artifacts in a variety of contexts (from sound design for film, to gaming, to sonic interaction design and sound branding) is not commonly adopted. As we will see later, efforts have been made independently in each of these disciplines to draft design processes that resemble to various degrees to the design thinking process shown in Fig.6. I believe that the lack of a shared methodology for approaching the design of sound is contributing to preventing the transition of sonification from a niche discipline to a widely applied tool for the representation of data.

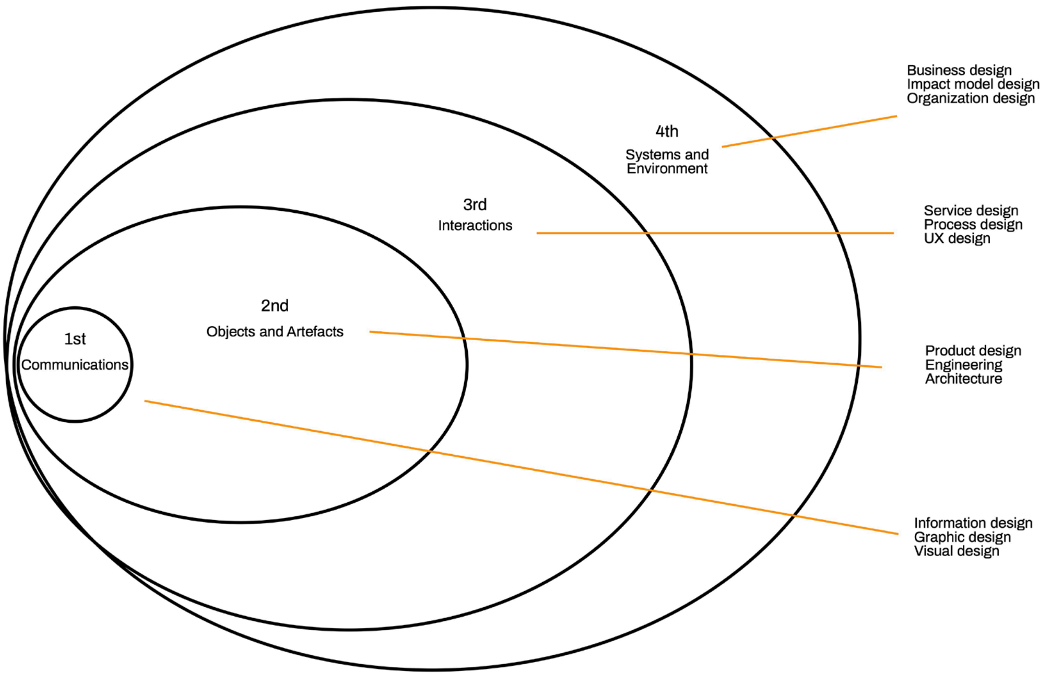


Fig. 7: The four orders of design according to Buchanan (cit.). Adapted from Martin C (2015).

Visual variables versus mapping problems- The issue of how specific values in a data set are mapped to specific acoustic dimensions and the choices behind this process have been raised and are sometimes referred to as the *mapping problem* (Roddy and Bridges 2016, p.67; Vickers and Barrass cit. p.153). Unlike in the academic field of visualization, where the so-called *visual variables*[8], which define and describe the “graphic dimensions across which a map or other visualization can be varied to encode information” (Roth 2015) are learnt by any freshman student, data sonification authors still have to address and defend almost on a case-by-case basis the choices they make during the design process in terms of how data values are translated into sound values.

8. The 12 visual variables in their current definitions are based on the work of the cartographer Jacques Bertin and were first described in his 1967 work ‘Semiology of Graphics’.

The answer to the question "How might abstract data be converted to sound such that the structure of the data is accessible to interpretation?" (Roddy and Bridges cit., p.67) is still far from being answered, despite some notable efforts that either tried to ground mapping choices within wider frameworks such as embodied knowledge (Roddy cit.) or attempted to systematize existing literature on a specific topic in search for recurring strategies that could be standardized (Dubus and Bresin 2014). Nonetheless, the mapping problem as it was formulated by Flowers (2005) i.e., that meaning does not arise naturally when data are represented with sound, questions the very first definition of data sonification and its assumptions: that in the process of sonification, *numerically represented relations* in a real-world phenomenon are translated into cognitively equivalent *relations in an acoustic domain*, and that therefore, contrarily to what Flowers believes, representing the contents of complex data sets with sonification will necessarily lead to the emergence of meaningful relationships in the data.

Over the past two decades, efforts to demonstrate such correlations between the characteristics of a real-world phenomenon, its numerical representations and corresponding characteristics in the auditory field has returned mixed results. In their 1996 seminal article 'Mappings and Metaphors in Auditory Display', Walker and Kramer were mainly "concerned with identifying situations where the choice of data-to-display mappings can have a significant effect on performance, especially for data dimensions that may be represented in a wide variety of auditory displays" in the hope that this investigation would lead "to the development of mapping guidelines applicable to auditory displays in a wide range of task domains" (Walker and Kramer 1996, p.407). Experimental results concerning the dimensions of temperature, pressure, size and rate actually concluded that "data-to-sound mappings that seem intuitive to a sound designer may actually result in less effective performance" (Ibid., p.411). In a famous experiment involving two groups of testers, one comprising visually impaired listeners and the other sighted listeners, Walker and Lane demonstrated that, in some cases, the two groups interpreted the same mapping in opposite directions. In the most striking case, the mapping frequency-dollars resulted in sighted listeners interpreting an increase of pitch in the sound as an increase in wealth whereas visually impaired listeners interpreted an increase in wealth with a decrease in pitch (Walker and Lane 2001).

Again, the underlying theoretical issue at stake seems to be the tension between a supposed objective and numerically quantifiable information contained in a data set and the struggle of representing such information through a sensory phenomenon (sound) that we interpret through the filter of perception which is, as such, exposed to the subjectivity of the individual listener. One of the objectives of this thesis is to frame the so-called mapping problem in terms of design decisions i.e. specific choices that authors have to take (and make explicit) in order to design sonifications that can really have a meaningful impact on our daily experience of data.

The lack of experimental validation- Walker and Kramer concluded their investigation by acknowledging that "it is crucial to empirically test an auditory display with listeners representative of the final users" to inform iterative prototyping and verification of sonic interface designs (cit., p.411). Twenty-five years later, the same need is still advocated within the data sonification community (Axon et al. 2017; Roddy and Furlong cit.; Barrass cit.; Vickers and Barrass cit.; Vogt 2011). Admittedly, a design process that engages final users from the beginning (when decisions on mapping strategies and sound material are taken) and ideally continues iteratively during prototyping and until final delivery, is far from being the norm in the data sonification community. Moreover, there is no shared experimental protocol to assess the appropriateness of a sonification to real-world situations, to the detriment of its chances of becoming a more widely used data representation method. Dubus and Bresin (cit.) showed that only 11.7% of the 60 projects they considered actually went through some kind of structured assessment procedure, which would help the field of sonification "gain maturity" (ibid., p.14).

Can design, as a structured *undisciplined* discipline (Bremner and Rodgers 2013) help sonification reflect on its unique *way of doing* (Cross 1982) and developing processes and tools to bridge the divide with other data representation methods? As Ken Friedman puts it, "The challenge of any evolving field is to bring tacit knowledge into articulate focus. This creates the ground of shared understanding that builds the field" (Friedman 2000, p.13). In the past decades, the field of design has struggled to develop flexible and adaptable experimental protocols able to do justice of the designerly way of doing and knowing.

It is to these protocols that I will turn in the upcoming section in order to explore the first two research questions: can we frame a design methodology to approach data sonification projects from prototyping to testing? How can we evaluate whether one sonification is better than another?

A role for (sound) design?

One of the main assumptions of this thesis is that in data-intensive societies - where the complexity of social issues is increasingly evident and inevitable and the visual channel is often overstimulated- sonification can expand the domain of action and "reconfigure [..] from an instrument solely for scientific enquiry into a mass medium for an audience with expectations of a functional and aesthetically satisfying experience" (Vickers and Barrass cit. p165). This transition will require, among other things, the definition of a specific set of knowledge, skills and methods that I believe design as a discipline can provide. A design-driven approach to the sonification of data for communicating with and engaging non-expert publics has been explicitly advocated within the research community. As Barrass puts it, "The design of a sonification requires a holistic approach to functionality and aesthetics that integrates art and science. A design approach accepts messy problems with multiple solutions whilst also allowing for critical reflection and empirical evaluation. Design knowledge is built through an iterative process of refinement of both aesthetics and functionality" (Barrass 2012 cit., p178) and as such, it could provide sonification with the required methodological infrastructure to bridge the gap between art and science.

Undeniably, there is still a long way to go for sonification to become a widely recognized medium for representing data. Once again, a parallel with visualization, and in particular the transition of data visualization from a scientific discipline to a mass medium that communicates to a lay, non – expert audience, could be useful to set the boundaries of a multi-faceted and complex theme. Reflecting on the potential role of design in guiding this transformation, authors such as Lau and Vande Moere (2007) and Manovich (2008) proposed the term *info-aesthetics* when calling for a wider contribution of design in information visualization. This contribution would inject often overlooked considerations of the aesthetic dimension of the visual experience and prompt a more intuitive understanding. The aesthetic dimension of design and its role in the ideation process and, more importantly, in the experience of the final user is to be found both in data visualization as an artistic expression and in commercial products created by professional designers as a parameter of 'attractiveness' i.e., appeal or beauty (Vande Moere and Purchase 2012). The attractiveness of a visualization "may compel the user to engage with the data, enabling more effective communication of the information itself" (ibid., p.363). Designing with attractiveness as well as functionality in mind concerns a process where choices are made based on a designerly way of knowing and doing which is, still today, mostly tacitly acquired and developed through practice. The explicit acknowledgment of the rationale behind design choices in data visualization is, according to Vande Moere, a much-needed step for the development of guidelines for the creation of successful (efficient and compelling) visualizations that intend to improve the engagement and meaning-making in the context of human-data relationships.

The correspondence between a functional and an aesthetic approach to sonification has been raised on various occasions within the data sonification and auditory display community. The lack of aesthetic considerations resulted in sonification being described as intrusive, distracting, causing listener fatigue, annoying, subject to issues of display resolution, precision and comprehensibility according to Vickers and Barrass (cit., p.154) and even "utterly ugly" according to Kramer (cit.).

For Roddy and Bridges (cit., p.67) "Care must be taken when considering aesthetics in a sonification context" while distinguishing between cosmetic practices which "simply aim to produce an attractive and easily listenable sonic result" and an aesthetic approach which aims, instead, "to frame and shape the qualities of the listeners' sonic experience as a means of communicating information about a data source". Furthermore, Leplâitre and McGregor (2004) empirically investigated the relationship between performance-based tasks and aesthetic judgement. Results seem to indicate that functional and aesthetic properties of an auditory display cannot be dealt with independently. Again, design as a discipline has a solid history of dealing with the complexity of the relationship between aesthetics and functionality, a well-known example being the work of Donald Norman on emotional design (Norman 2004).

Within the landscape of sound related practices, we find the same tension in the definition of sound design as “the ambition to design and explore the creation of new sounds that have a communicative function which is supported by aesthetic qualities and form.” (Pauletto 2014).

In the first stage of my doctoral research, I hypothesized the existence of a role for sound design in the definition of a designerly approach to data sonification. This hypothesis forms the basis for the first research question “What is the role of the (sound) designer in data sonification? Can a designerly approach to sonification make the difference in creating better (more efficient and engaging) representations of data?”. In the following paragraphs I will expand on these questions taking a closer look at the field of sound design and existing sound-related design processes.

RQ1: What is sound design- According to Susini and colleagues, the term sound design as we intend it today emerged over the course of the 20thcentury as the convergence of two main fields, sound design for film and the concern around the quality of sound in everyday products (Susini, Houix, Misdariis 2014). Walter Murch, who is believed to have coined the term ‘sound designer’ to define his own role in 1970s Hollywood movie productions, defines it as the power to capture “ordinary sounds and reorganize them” thanks to the magnetic tape (Murch 1992, p.xv). The idea of abstracting concrete, everyday sounds from their primary context of experience to be used as the raw material of new auditory experiences can be traced back to the work of Pierre Schaffer at Radio France in the 1940s. His concept of a musique concrète - a new music that transformed, thanks to recording and sound diffusion technologies, an everyday listening experience into a new art form - has had a long-lasting influence on both contemporary electronic music and sound design as we know it today.

An increased awareness of ordinary, everyday sounds also emerged during the course of the 1970s when R. Murray Schafer (1977) brought to the attention of the public the concept of soundscape (the collection of everyday sounds we live immersed in), key sound (the background general key of an environment) and sound mark (a sound which is unique to an area). In inaugurating the interdisciplinary research field called 'acoustic ecology' (Truax 1978), Murray Schafer was emphasizing the need to actively preserve a natural soundscape that is constantly threatened by the anthropic, 'schizophonic'[9] sounds increasingly produced by urbanized societies. During the course of 1970s and 1980s, along with the advent of noise level regulations in the public space and the emergence of a new sensitivity toward an ecological approach to the environment, the idea that intruding, anthropic sounds had to be limited as much as possible even in everyday products started to be adopted by industry (specifically, car and home appliances manufacturers) thus introducing the concept of sound quality in product design. Still, though, the main approach was to limit the impact of unintentional sound on the everyday soundscape. From the 1990s, advancements in the field of both audio technology (with affordable miniaturized sound diffusion systems reaching the market) and digital audio production (with sophisticated sound synthesis programs made available to both professionals and amateurs) led to the emergence of an approach that, instead of trying to limit the impact of unintentionally generated 'bad sounds' proposed to intentionally design 'good sounds' that addressed specific users and were experienced in specific contexts. Today, beyond the boundaries of a "relatively small subset of sounds created for digital media" (Susini, Houix, Misdariis cit., p105) the design of sounds is applied to any context in which new sounds are intentionally created through a number of different processes (recording the environment and manipulating everyday objects as in foley or produced through sound synthesis). These new sounds are defined both by their function and their form i.e., by their informative content and by their aesthetic qualities and are applied in a variety of fields from sound design for film to sonic interaction design and sound branding.

9.
Schizophonic sounds are sounds that have been abstracted from their original contexts through recording and are reproduced in the environment through electroacoustic means. Interestingly, this definition (in negative terms) closely echoes the definition by Pierre Schaffer (in positive terms) of musique concrete as a new liberated artform.

Sound design for film

Perhaps the most well-known field of application of sound design, it benefits from a long-established artisanal tradition now condensed in numerous educational curricula at both under- and postgraduate level. Tacit knowledge on how to create new sounds with the goal of supporting the narrative and triggering emotions in the audience has been condensed to form new generations of sound designers and shared over the years through successful publications and resources such as ‘Sound Design’ by David Sonnenschein (2001). The community of sound designers for film and the gaming industry is a particularly active one, with numerous online forums, portals and groups where professionals and amateurs share experiences, tips and sound libraries on a daily basis. Moreover, authors like Michel Chion have contributed to the development of a theory on the role of sound in shaping the experience of the public in the art of cinema and musicals. In ‘The Audivision’ (cit.), he presents a rather sophisticated theoretical model that has been very influential in the theory of sound design

The model attempts to describe how sound adds value to the visual channel in building the sense of time and space and the narrative of a movie. Far from being a mere embellishment of the visual experience, sound adds to the image an “expressive and informative value with which it enriches a given image so as to create the definite impression, in the immediate or remembered experience one has of it, that this information or expression ‘naturally’ comes from what is seen, and is already contained in the image itself” (Chion cit., p.5). This audio-visual illusion is intentionally triggered in the listener by designing sound according to a specific framework which Chion defines by analyzing numerous cases borrowed from the history of cinema. A particular relevance is given to how the public perceives sound in the context of the cinema experience, in what Chion defines ‘listening modes’. I will describe the three listening modes identified by Chion – causal, semantic and reduced - in greater detail in the following chapter. It is worth noting here that the attention to the audience’s listening experience is of great importance in the field of sound design for film as it can help define how certain auditory stimuli can trigger a specific emotional response, thus providing a design framework that can bridge the gap between the emotions the film director wants to evoke and the way the public perceives these emotions aurally (Hillmann and Pauletto 2014).

Sound design for video games

In game design, sound is used to both inform the player on his status (for instance on motion – through the sound of steps and health – through breathing) and on the status of the game’s world (for example, on the weather or the time of the day or certain characteristics of indoor spaces). Additionally, and like in sound design for film, sound is used to trigger in the player an emotional response. Unlike sound for cinema, though, sound design for gaming has to take into account the principal elements of the game narrative: interactivity and non-linearity (Collins 2008), which represent a challenge mostly for the composition of game music. With the expansion in Virtual Reality technologies and the increased availability on the consumer market of integrated devices that allow the player to see and hear in 3D, sound design for gaming is reaching a new level in shaping the user experience of players. In particular, there is currently a focus on designing interactive localized sounds with enhanced spatial attributes to improve the sense of immersion in a realistic space (Murphy and Neff 2011).

Sonic Interaction Design

The way in which sound can be used to "convey information, meaning, and aesthetic and emotional qualities" (Franinović and Serafin 2013, p.vii) in any interactive context well beyond the specificities of codified multimedia experiences (such as video gaming) is the field of action of Sonic Interaction Design (SID). In design disciplines, according to Franinović and Serafin (ibid.), "sound has been a neglected medium, with designers rarely aware of the extent to which sound can change the overall user experience". The highly interdisciplinary field of SID (usually considered a branch of Human-Computer Interaction) "emerged from the desire to challenge these prevalent design approaches by considering sound as an active medium that can enable novel phenomenological and social experiences with and through interactive technology" (ibid., p.viii). It is worth noting that auditory display and sonification do have a space in the SID field when they are tackled from the perspective of interaction between the user and the sonification artifact (Franinoviç and Serafin, ibid.).

Sonic Information Design

Closer to the topic of this dissertation, Sonic Information Design is a term coined by Jeon, Walker and Barrass (2019) to frame a new field which would apply design research to data sonification, auditory display, auditory user interfaces and so on. Their proposal of a discipline explicitly founded on the design research paradigm, that "assumes that there is no universal or optimal solution and that different designs will be more effective for different users, tasks, and contexts" (ibid., p.2) certainly resonates with the assumptions, goals and approach of this work.

The focus on the intentionality in the design of sounds to trigger specific (emotional and informative) experiences in a specific context paves the way for the definition of a work process for the design of sound in all the contexts we presented in this paragraph. According to Susini and colleagues (cit.), a sound design framework is meant to answer specific questions on listening strategies ("How can the listening strategy be taken into account for the design of new sounds?"), types and function of sounds ("What types of sound analogy best promote interactions with an object or an interface?") and information-to-sound mapping ("Can sound help to learn and to control different simultaneous streams of information?"). Such preliminary questions "help sound designers make relevant choices instead of starting from scratch." (Susini, Houix, Misdariis cit., p.114). As we will see in the following chapter, very similar questions emerged during the interviews with authors of sonifications and helped shape the sonification canvas as a tool for the designer to gain an awareness of her decisional process during the production of a new sonification. The following paragraph reviews existing processes for the design of sound in contexts such as audio branding, sound design for film and multimedia and auditory display with the goal of laying the foundations for a design-driven approach to the production of sonification projects.

RQ1: Drafting a designerly way for sonification- According to Jorge Frascara, "Visual communication design, seen as an activity, is the action of conceiving, programming, projecting, and realizing visual communications that are usually produced through industrial means and are aimed at broadcasting specific messages to specific sectors of the public." (Frascara 2004 cit., p.2). Vande Moere and Purchase (cit., p.364) echo this description and apply it to the specific field of data visualization, highlighting that "Key decisions regarding intent and visual properties are made as part of the design process; they contribute to the attractiveness of the visualization and are motivated by design rationale – whether explicitly stated or not". The exercise of making explicit a mainly implicit design rationale is far more common in the field of design (both in research and in industry) than it currently is in the field of sound design. A recent exception is the effort by the Sound Design and Perception Team at IRCAM to define a work process which is largely inspired by “classic design processes that are usually proposed when designers work on a project” (Susini, Houix, Misdariis p109). Figure 8 illustrates their proposal which roughly covers three main phases: a preliminary research phase (Analyzing), a creative phase (Composing) which leads to the formalization of a prototype; and a testing phase, in which the prototype is tested with users.

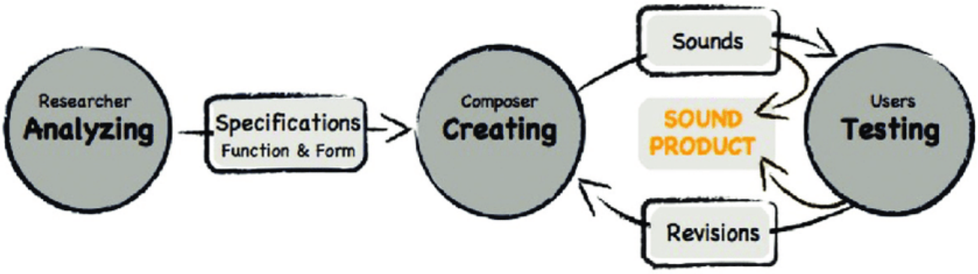


Fig. 8: The sound design process proposed by the Sound Design and Perception team of IRCAM (Susini et al. cit.).

Similar design processes can be found in the commercial field of sound branding (or audio branding). Sound branding deals with the design of sounds in support of brand communication. Sound identities, sound logos or complete sound suites that integrate sonic interaction design for physical products, sound and music for advertising and background soundscapes for physical spaces are produced in collaboration with the brand identity team to foster brand engagement and support marketing efforts.

Sound branding is a growing field and its methodology is assessed through publications like ‘The Sonic Boom’ (Beckerman and Gray 2014) and in industrial arenas like the annual Audio Branding Congress[10] organized by the Audio Branding Academy (ABA). ABA also confers industry awards where quality parameters are discussed and established by peers and it publishes an annual yearbook (Bronner Ringe Hirt 2020) where professionals reflect on their work and share the design rationale behind their projects. An example of a workflow that echoes the traditional design process is the one shared by the sound branding agency GROVES – Getting your brand heard[11] (Figure 9). Similarly to the sound design process proposed by IRCAM, a preliminary explorative phase is followed by a concept development

10. https://www.international-sound-awards.com/

11. https://groves.de/en/

phase and a design phase. User testing is replaced by the definition of brand sound guidelines for the client and by sound tracking, an activity which is meant to measure the impact of the sound branding initiative against metrics such as brand engagement and revenues.

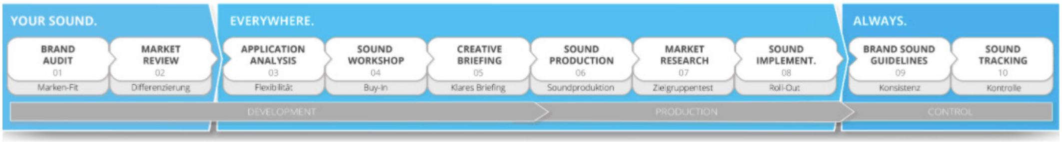


Fig. 9: GROVES sound branding process. Copyright: https://groves.de/en/

In sound design for film, Sonnenschein (cit.) proposes the sound map as an iterative support for the sound designer. For each sequence of the movie, references to the sound elements that have to be designed (voice, music, sound effects and so on) are placed on a horizontal axis, with the vertical axis representing the movie timeline. As the movie production proceeds, the sound map is filled with details on specific sound elements, foley, music themes and so on. Additionally as shown in Figure 10, a visual map of the movie's narrative with a graphical representation of the weight of the soundtrack against the different phases of the narrative can be sketched as a guide to the overall sound design effort (Bishop and Sonnenschein 2012).

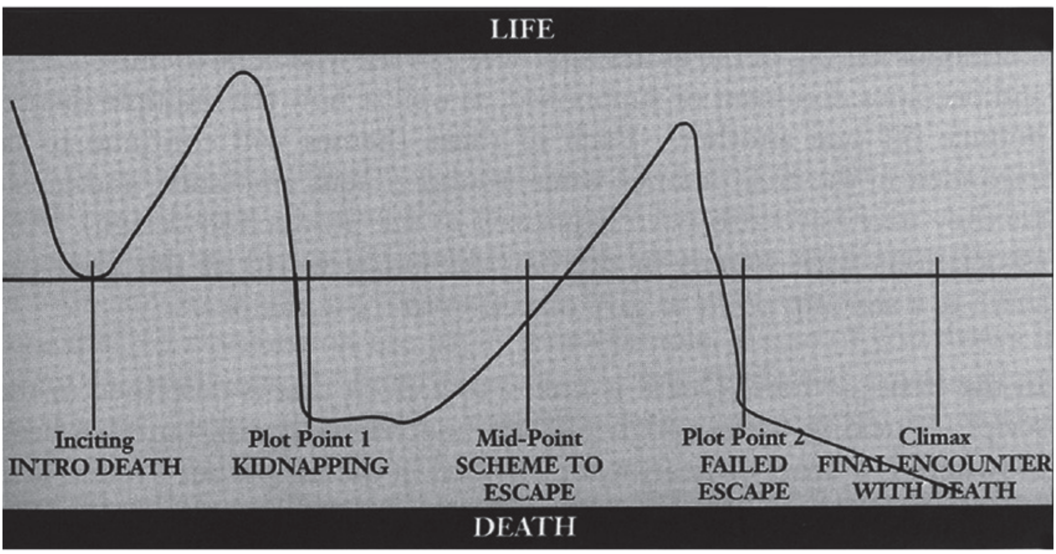


Fig. 10: Sound map for sound design for film. Bishop and Sonnenschein 2012.

In auditory display research, Brazil and Fernström (2009) defined a design framework which is divided in two phases: sound creation and sound analysis. The sound creation phase echoes some of the traditional steps of the design process in that it identifies phases such as:

- Context and auditory display definition.
- Selection of sounds.
- Creation of the sounds.
- Listening to the sounds as a form of self-evaluation and process iteration.

In the field of data sonification, de Campo proposed a data sonification design space map (de Campo 2007) as a support to systematically approach the choice of sound based on the type of dataset (ibid., p.1).

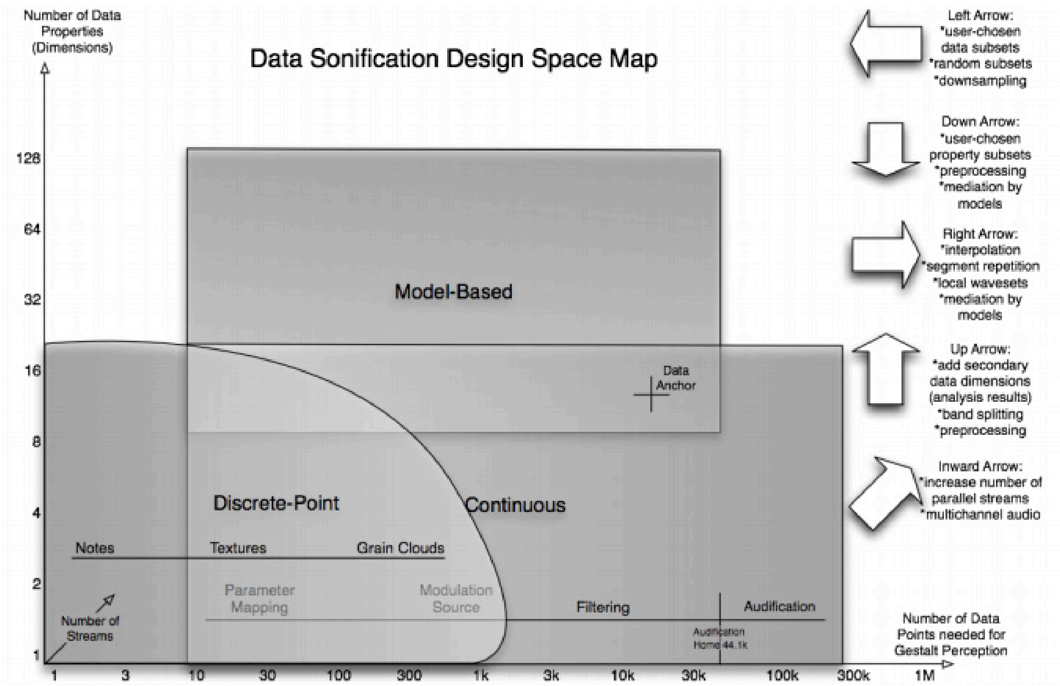


Fig. 11: Data sonification design space map. de Campo 2007.

The map in Figure 11 does not seem to refer to a traditional design process as described by both the IRCAM Sound Design Team and sound branding professionals (and, to some extent, by sound design for film producers). Rather, de Campo's map is a structured system to guide mapping strategies when dealing with specific datasets and specific audio parameters independently from the context of usage or the communicative goal of the project.

Barrass (2003) described several sonification design patters inspired by Christopher Alexander's architectural and urban design patterns (Alexander, Silverstein, Ishikawa 1977) as a way to catalogue good solutions to common sonification problems in a given context. The design patterns take the following form (Barrass 2003 cit. p172):

> **"IF you find yourself designing** sonifications for
> **small, mobile, info-tech consumer products**
> [...]
> **with the problem that**
> the sounds in most products are either alarms or embellishments, and most people are not familiar with sonifications that convey other types of information
> **THEN**
> Design a sonification that provides information that makes the product more useful, usable and enjoyable"
> and have the goal of helping the field of sonification develop a technical shared vocabulary that would foster the development of the field.

Despite the lack of a unified strategy, the last two efforts denote the interest of some authors in the definition of a design-driven process for the design of sonifications. The following section looks at how the community has tackled the issue of the experimental validation of sonification projects.

RQ2: Experimental validation: preparing sonification for the real-world- The lack of a standardized procedure for designing and experimentally validating potential real-world applications of sonification has been pointed out recently (Hildebrandt and Rinderle-Ma 2015; Axon et al. 2017 cit.; Lenzi et al. cit.) but is not a new concern. In 2011, the 'Sonification Handbook' urged researchers in the field of auditory display to familiarize themselves with the robust experimental protocols described in detail in Chapter 6 of the Handbook (Bonebright and Flowers 2011) in order to validate the performance of their sonifications. Issues of usefulness, effectiveness and efficiency of sonification as an independent modality to visualization have confronted researchers on several occasions (Walker and Nees 2011 p14; Kramer et al. cit., p.19). Additionally, a recent review of 60 sonification projects by Dubus and Bresin (cit.) found that only about 11% of the resulting prototypes had been validated with users.

But what would an experimental protocol for data sonification projects look like? In the previous paragraph, Susini and colleagues highlight the need for sound design prototypes to be tested with users. Specifically, they propose using listening tests, preference maps and methodologies borrowed from experimental psychology "until a prototype sound is obtained that fulfils the perceptive expectations in terms of function (or aesthetics)." (cit., p.110). Brazil and Fernström (cit.) propose using qualitative methods such as sonic maps and ear-witness accounts to evaluate the subjective experience of sound in auditory display as part of the design framework presented in the previous paragraph. In the specific field of data sonification, Vickers and Barrass (cit., p.161) suggest the usage of the psychologically based qualitative methodology IPA (interpretative phenomenological analysis) on the grounds that "Traditional metric- and task-performance based techniques have been used to measure sonification design factors such as accuracy, recall, precision, efficiency, etc. Whilst one could measure the improvement on performance of auditory displays that have been designed to maximize their aesthetics, aesthetic judgment itself remains primarily experiential and so we can envisage using qualitative tools like IPA not only to gain more understanding of how users experience sonifications, but to evaluate the aesthetic dimension more richly." Barrass (2016) used annotated portfolios while experimenting with sonification artifacts (i.e. sonifications embodied in acoustic tridimensional objects, in this case, a singing bowl). Annotated portfolios are a design research method developed by Gaver and Bowlers (2012) in which images of design artifacts are annotated with labels that later help identify design dimensions that can usefully address broader concerns in the research community. As we will see in the second part of this thesis, I tried to identify an experimental protocol that integrates quantitative and qualitative research for the assessment of sonification protypes in a real environment and with real users, through the deployment of the two Design Actions.

This chapter presented results of a preliminary research that was conducted with the goal of expanding on the assumptions of the research questions and in particular, on the definition of sound design and the evaluation of existing design and validation processes in sound-related disciplines. The following chapters will apply the concepts that emerged in this first phase to the development of processes, protocols and tools for a design-driven approach to data sonification.

CHAPTER 2 —
Exploration: choosing the key

Soundtrack: *M. Ravel, Piano Concerto in D minor.*
Arturo Benedetti Michelangeli.

Intentionality as a pre-requisite for design

According to Frascara, communication design "has an impact on the public's knowledge, attitudes, or behavior in an intended direction." (Frascara cit.). As we saw in the previous chapter, the Sound Design Team at IRCAM also defines sound design as the creation and manipulation of sounds with the intention to communicate information (Susini Hoiux and Misdariis cit.), thus echoing Frascara's definition. Intentionality, or the intentional creation of a new artifact with the goal of communicating a specific message to a specific public, seems to be a pre-requisite of all design practices, be it through the medium of image or of sound. In the field of data sonification, as the needs of publics are increasingly being taken into account, the role of authors' intentions also needs to be addressed (Barrass 2012 cit., p.178), thus paving the way for a design-oriented approach to the communication of data. Whether as a tool for scientific analysis or as a communicative experience, data sonification involves different degrees of intentionality, i.e., deliberate decisions to address specific needs, in a given context and with a purpose, when transforming data into sound. However, for a practice still in its infancy it is unclear whether the authors of sonifications are ready to manage – and take responsibility for – the inevitable communicative ramifications of their productions and their agency: to what extent are they aware of it? Are they taking into consideration and intentionally addressing that responsibility – as a design approach would require? In order to define a design-driven data sonification space, I explored the notion of intentionality as a pre-requisite for a designerly approach to sonification through the discussion of five recent cases that aim at engaging publics with social issues. The selected cases are distributed on a scale of intentionality based on the explicit statements of their authors while describing the process and the goals behind the sonification. From the highest degree of intentionality of Egypt Building Collapses by the activist group Tactical Technology Collective, to the use of sonification in data journalism and sound art on sensitive social issues, intentionality has been considered an indicator of a designerly approach that can position sonification as a meaning-making medium in its own right.

Unpleasant sounds and causal listening: sound as a connecting element- The Tactical Technology Collective[12], together with the Shadow Ministry of Housing and the Egyptian Initiative for Personal Rights, explored sonification in the 2013 project Egypt Building Collapses. Data on one year of accidents involving the sudden collapse of residential buildings in Egypt are visualized on the website to raise awareness of a serious issue affecting Egyptian society, that resulted in 192 casualties and more than 800 homeless families in only one year. Collapses are mainly due to poor public policy and regulation of private buildings' construction and bad planning (Sobliye and Mortada cit., p.208). The authors used sound as a "connecting element between the real experience of the collapse of a building in Egypt and the figures that describe it over time" (Briones 2018). Sound becomes a complement of the visualization to emphasize the impact of the issue on the lives of real people and to increase awareness about the real tragedies behind numbers of building collapses.

12. https://tacticaltech.org/#

Data that illustrate the occurrences of such accidents for one year are used to build a 2:35 minute soundscape that is streamed on the project's website to accompany the visual experience. Upon accessing the website, a window suggests that visitors wear headphones or increase their speakers' volume, thus alerting them to the sound they are about to experience (Figure 12).

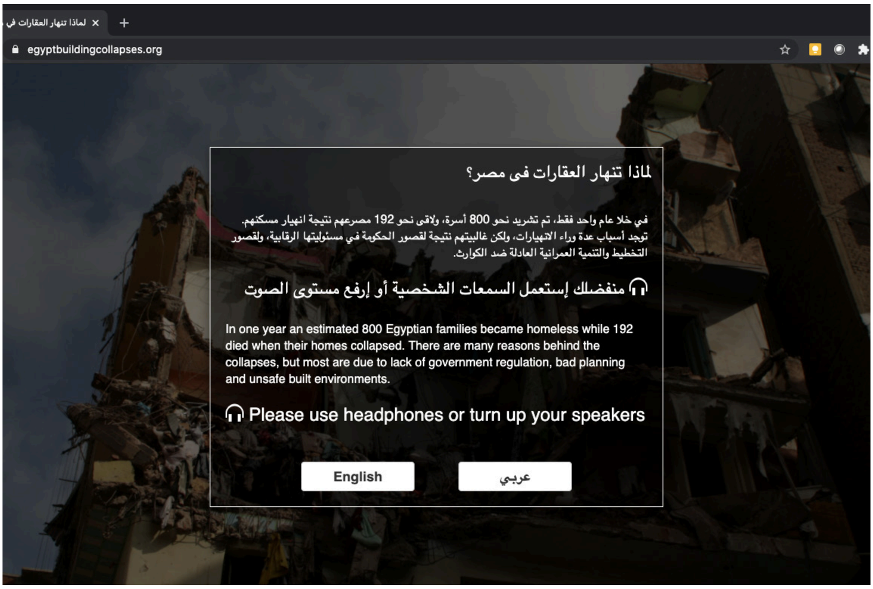


Fig. 12: Egypt building collapsing website homepage which alerts visitors about the sonification experience

Figure 13 illustrates how the soundscape is broadcast on top of the website's main page together with the visualization of data, which shows the number of building collapses along with the number of injured, dead and affected families. Interacting with the visualization, visitors can access further information detailing the reasons behind injuries, deaths and family displacements, as well as zoom in on a specific area of Egypt. As for the sound experience, the soundscape is composed by sonifying each accident with a single sound of a collapsing building; the density of the sound samples gets higher over time as accident rates increase. The authors considered various options for the sound: "The first was calmer and softer; it could be described as 'relaxing' or even 'meditative'. The second test featured more literal sounds of falling bricks and unstable foundations" (Sobliye and Mortada cit., p.215). Despite being conscious of the potentially intruding and annoying nature of sounds on the internet, particularly rough sounds such as building collapses, the authors chose the second option in order to "tell stories of social injustice" that may even be intrusive and cause annoyance when confronted.

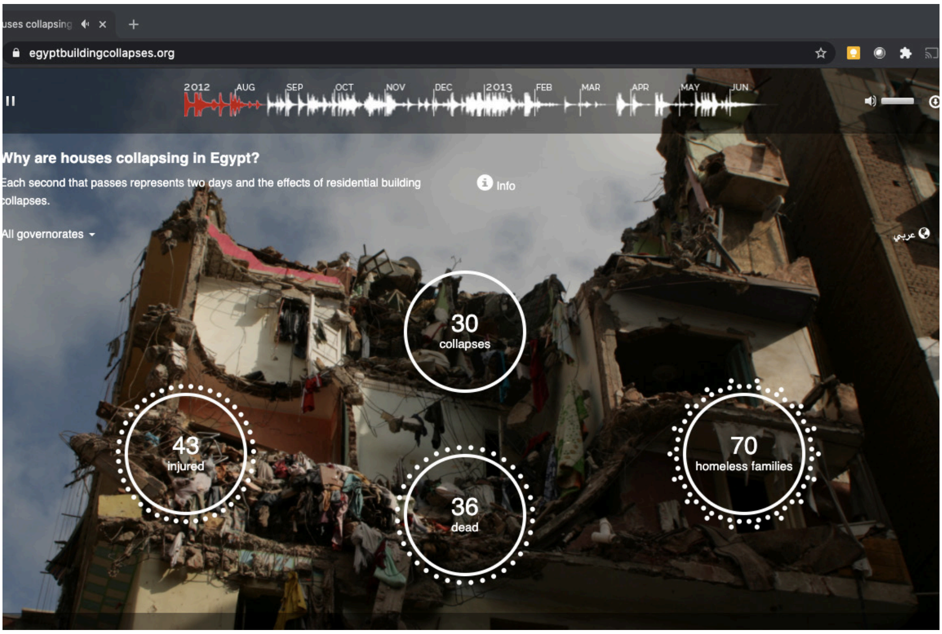


Fig. 13: Data are visualized and sonified.

The authors considered various options for the sound: “The first was calmer and softer; it could be described as ‘relaxing’ or even ‘meditative’. The second test featured more literal sounds of falling bricks and unstable foundations” (Sobliye and Mortada cit., p.215). Despite being conscious of the potentially intruding and annoying nature of sounds on the internet, particularly rough sounds such as building collapses, the authors chose the second option in order to “tell stories of social injustice” that may even be intrusive and cause annoyance when confronted.

The sonification of Egypt building collapses is a good example of what Chion calls causal listening (Chion 2015, p.312): as a result of the choice of sound material, the listener will tend to associate the sound directly with its source i.e., with the event that caused the sound itself (in this case, the collapse of a building). Chion notices that the use of indexical sounds (sounds in which the connection with the phenomenon producing the sound is not arbitrary but rather connected to the phenomenon itself (Vickers and Hogg 2006, p.213) is made possible by a shared cultural experience such as that which happens with cinema. As listeners, it is through the experience of cinema (or, more recently, gaming) that sounds caused by events can be recognized and correctly interpreted despite never having been experienced first-hand (such as the collapse of a building, a gunshot, or an extraterrestrial language) (Chion 2015 cit., p.319).

Ominous sounds and sonic memorials: sonification for on-air data journalism- The American award-winning Center for Investigative Reporting[13] has been experimenting with data sonification since 2015 through its online platform Reveal, which also hosts a radio show. In its first production, 'The Oklahoma Shakes'[14] , a soundscape was produced that connects data on earthquakes occurring in Oklahoma from 2004 onwards, with the goal of providing insights on the exponential increase of such events (allegedly due to the increase in the oil and gas industry exploitation of Oklahoma land), and this sonification was later broadcast on the radio. The timeline of the soundscape is obtained from the chronology of earthquakes over the years, and each sound in the timeline is mapped to data on each single earthquake: the bigger the earthquake, the louder and more low-pitched the sound. In this first foray into the sonification world, the design of individual sounds (a bell-like synth sound obtained via MIDI processing) is based on minor chords, a more 'ominous' key than major. In the words of the author Michael Cory (cit.), "choosing an ominous key – or any key – is as much an editorial decision as choosing colors on a map visualization. In this case, there's not really any way around the fact that suddenly experiencing hundreds of new earthquakes every year is an ominous development, so I felt justified in the decision". Here a cultural interpretation of musical and emotional values is intentionally used by the author to determine design decisions on how to map data to sound, loosely suggesting to the listener a connection between 'sad' sounds and sad events such as earthquakes. It is unclear, though, how much we can rely on cultural clichés – e.g., that minor keys are considered, in Western culture, to carry a sad message – as the design of an experience that is broadcast on the radio should be universally understood by any listener without prior training.

13. https://revealnews.org/

14. https://www.revealnews.org/article/listen-to-the-music-of-seismic-activity-in-oklahoma/

Another example, 'A sonic memorial to the victims at Orlando's Pulse nightclub'[15] takes a different approach that seeks to design sounds that bear a specific message and explicitly communicate the intention of the authors as explained by Jim Briggs (2016). On the 12th of June 2016, an armed gunman entered the Orlando nightclub Pulse, in Florida, United States and opened fire on the guests, killing 49 people

15. https://revealnews.org/blog/a-sonic-memorial-to-the-victims-at-orlandos-pulse-nightclub

16. https://en.wikipedia.org/wiki/Orlando_nightclub_shooting

and injuring 53 others[16]. The soundscape is built using the birth year of each of the 49 victims of the shooting, from 1966 to 1998, when the youngest victim was born. Each person is represented with a bell-like tone, chosen because of the iconic value of bell sounds across different cultures, “from European village bells and the way that they signify important moments in communities to gamelan performances and the additive effect of overtones produced by a body of many different, interdependent instruments” (Briggs cit.). From the oldest to the youngest victim, from a lower to higher pitch, ‘A sonic memorial’ composes a sonification that is meant to be “funereal, but also celebratory”. Over the course of the sonification, the different tones (the different persons’ birthdates) accidentally cross to create melodies which match, for better or worse, just as people match better with some individuals as opposed to others throughout the course of their lives. Through a poetic interpretation of the simplest dataset, the authors created a sonic representation which powerfully conveys that “there is beauty in those cycles and new life within every beat” (Briggs cit.) until the soundscape is abruptly interrupted at the year of the shooting. Designed to be broadcast as a radio show and as a sort of ‘threnody’ to the victims, the project aims at communicating the value of life and the tragedy of mass-shootings in the United States. In order to do so, the authors describe how data were mapped to sound dimensions by building upon socio-cultural values shared by both the designers and the listeners of the sonification. The choice of musical instruments, the sequence of the sound events over time, and the decision to use tuned (musical) sounds, which would occasionally create a sense of harmony in the composition, are explicit choices. The authors appear to rely on shared metaphors (Johnson 2007; Roddy and Furlong 2015, p.183) to emotionally engage publics who do not need specific prior musical training.

Experiencing forty years of global migrations through sound: generative art and sonification- A different approach is taken by the sound artist Brian Foo, also known as the Data-Driven DJ. Foo is the author of a number of sonification projects in which generative art (using data as a compositional tool to create art experiences) and social responsibility (using sonification to engage publics in ways that resonate with events represented by the data) are mixed to different degrees. In ‘Distance from home’[17], the United Nations global refugee dataset from 1975 to 2012 is used to generate a song in which different audio parameters are mapped to different dimensions in the data. Specifically, 'The quantity, length, and pitch of the song's instruments are controlled by the volume of refugee movement and distance traveled between their countries of origin and asylum' (Foo 2012). Unlike the two previous projects, Foo approaches the sonification process by focusing on the design of a customized software rather than on the meaning of the content and the listening experience. The software determines the creation of a musical composition where the structure changes over time following the behavior of the dataset. In his words, Foo's goal is to help the listener to "intuitively and viscerally experience the sheer volume of displaced populations and the distance they travel from their home country" (Foo, cit.). As the years go by, the volume of instruments playing and the duration of the notes they play increases as the global pitch of the song lowers: it means more people are leaving their homes to travel longer distances.

17. https://datadrivendj.com/tracks/refugees/

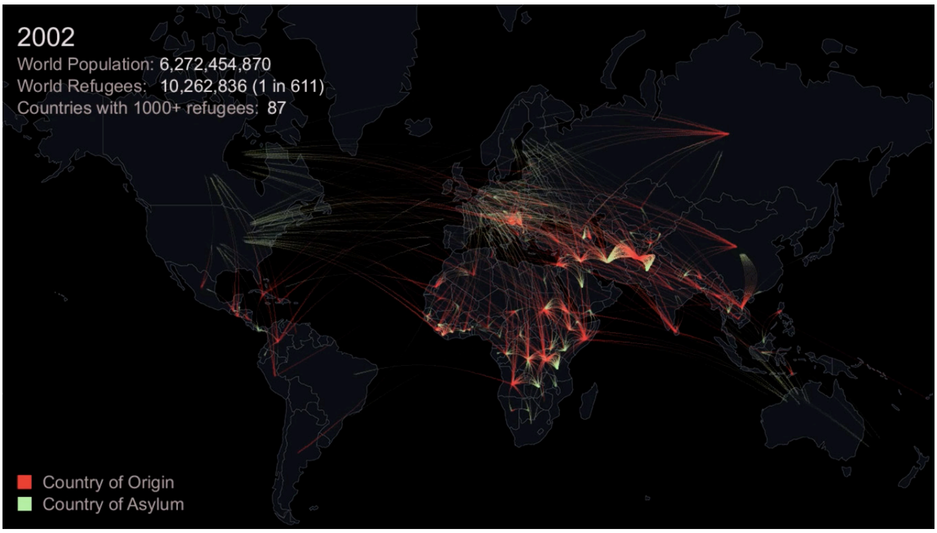


Fig. 14: Frame from the video animation for 'Distance from home'.

The experience of the sonification can be accompanied by visuals (Figure 14 illustrates a frame of the visualization for the year 2000) or only through sound. A link to the sonification without visuals is also provided for listeners who "prefer no visuals", perhaps an implicit recognition that when accompanied by visuals the nature of the listening experience would change. Conveying through sound any additional contextual information (which the artist defines as "the reason for displacement") is explicitly excluded by Foo in order to avoid "too much complexity". In this way, Foo's intention seems to be to elicit in the listener a semantic/figurative listening attitude (Chion 2015 cit., p.323) which requires prior training in order to infer the meaning of data from the different auditory dimensions of the composition. We could argue that, despite the choice of sound content being inspired by folk music as a cross-cultural expression of nostalgia and longing for home (Foo cit.), the usage of abstract musical dimensions still requires a higher cognitive effort in the listener to contextualize – and hence, make sense of – the data as compared to, for example, 'Egypt Building Collapses'.

Income inequality and agnostic sound material: the limits of sonification- In Two Trains, there is a greater reliance on musical dimensions to represent values in a data set without the representation of the context in which data are sourced. The project builds on the issue of income inequality among the districts of the city of New York, where unequal distribution of wealth is a larger problem than it is in other cities (Foo 2015), to build "a song with some exciting ups and downs" and at the same time "related with a topic that is relevant and current" (ibid.). Foo selected data on average income in the districts (Figure 15) crossed by subway line number 2 (Brooklyn, Manhattan, Bronx) as it looked like "the perfect song composition, with a build-up, climax and falling action." Average income values are mapped onto specific acoustic dimensions (pitch, rhythm, timbre) to obtain a sophisticated music composition which represents the change in income along places crossed by the two train lines.

Fig. 15: Frame from the video animation for 'Distance from home'.

Foo "scoured through the Internet" (Foo 2015 cit.) to identify songs and musicians that were representative of the NYC music scene in order to compile the sounds to be used in the sonification and in this way also pay homage to the music scene. It is worth noting that, based on this choice, the sound samples used for the sonification do not bear any relationship with the phenomenon represented by the data set. This is one way in which this approach is different from Foo's 'Distance from home' and the 'Egypt Building Collapses' projects. Foo explains that he "tried to select agnostic sound traits" (e.g. volume, dynamics) to correlate to median income rather than biased ones (e.g. sad vs. happy sounds, vibrant vs dull sounds) to further let the data "speak for itself" (ibid).

But can data really speak 'for themselves'? What is the role of the author in shaping the message that data convey to listeners? Additionally, and more specifically when considering the use of sound to represent and make data meaningful, to what extent can listeners unpack individual acoustic dimensions (such as timbre, pitch and rhythm) let alone associate these with specific data values? Can the increasing density of music in the poorer districts of NYC be interpreted as intended? For example, could the very same music be perceived as happy to some ears even in the poorest districts of NYC? Is the author's goal of being "as objective as possible" sufficient for accepting the responsibility that a communication process entails with a listener that, in this case, may receive a message that is the opposite of what is intended?

The analysis of the five cases helped me frame data sonification within a design process while introducing the criterion of intentionality as a condition. As can be seen (Figure 16), the highest degree of intentionality can be found in the 'Egypt Building Collapses' project, where sound is designed to explicitly help the listener to intuitively and emotionally connect with the sensitive social issue at stake. At the opposite end of the continuum – the lowest level of intentionality, as shown in Figure 16 – the author of 'Two Trains' declares his agnosticism to the context by aiming at being neutral: "I wanted to make a pleasant/exciting-sounding song so it could be palatable for the casual listener and experienced independently from the topic of income inequality. I didn't want to be accused of favoring one area of the city or the other because one sounds 'prettier' than the other" (Foo 2015 cit.).

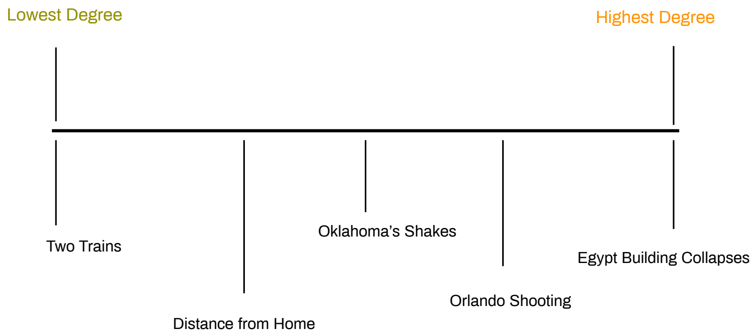


Fig. 16: Intentionality in Data Sonification.

Along the continuum, 'A sonic memorial to the victims at Orlando's Pulse nightclub' explicitly chooses a specific sound material (bells and bell-like tones) that will help listeners emotionally engage with the event represented by data, whereas in 'Oklahoma's shakes' a similar intention is still embryonic. In 'Distance from home', the absence of an explicit relationship between the design of sound and the meaning of data results in a rather abstract listening experience and prior training might be required to garner insights from the dataset.

This first evaluation of projects based on their authors' intention reveals that sonification seems to still be frequently guided by technicalities in the transfer and mapping of data into the language of sound (what we previously called the mapping problem) instead of a critical consideration of the effects and implications of this translation process for the listener. The lack of context, a reductive approach to the representation of the phenomenon behind the data, and the presumption of objectivity in data often prevent an effective integration of data sonification in the broader process of designing experiences that aim at reaching, engaging and informing, especially when addressing a wider, non-expert audience. The continuum of intentionality presented in this section could help authors increase awareness on the communicative needs sonification needs to address in its transition towards a medium of communication.

"Ritorno alle cose": the daily practice of data sonification

Through an analysis of five recent cases of data sonification for communicating social issues, I explored and defined *intentionality* as a necessary criterion to position the debate around the nature and the value of data sonification within the field of design – and specifically, communication design.

To recap, intentionality reflects the process by which a designer takes explicit decisions to address specific needs of a specific user in a specific context. Intentionality is a necessary criteria for design, but is it also sufficient, in order to fully discriminate and properly describe a designerly approach to sonification? What specific design decisions does the author of a sonification project need to take? What is the specific raw material that she has to explore and choose? What are the difficulties that sonification designers face, and what is the process that leads them to a solution?

In the previous chapters, we defined three areas of investigation related to the research questions, to outline a design space for sonification:

- The role of sound design.
- The definition of a design process.
- The definition of a validation protocol.

The intentionality criterion is not sufficient for defining all these three areas. I therefore decided to directly engage in a conversation with authors of recent sonification projects in the spirit of a neo-phenomenological return to things, an expression the Italian philosopher Luciano Anceschi (1998, p.78) uses in his lessons of aesthetics to suggest a renewed attention to the materiality of the arts. A 'return to things' is, in the words of Anceschi, the trigger of a reflection on the doing of art that tackles the philosophical, systemic question 'what is art?' (in Anceschi's specific case, 'what is poetry?') from the angle of 'how do you do art?'. A phenomenological approach to the doing instead of to the being is used to later inform the theoretical framework that intends to explain the meaning of art at a systemic level.

Seven authors, six men and a woman, with different backgrounds and geographical distribution, were interviewed over the course of several weeks. For each of them, I previously analyzed at least one specific sonification production which would loosely form the basis for discussion during the interview. Some of the cases had been included in the intentionality continuum. The format of the interview was semi-structured beforehand around topics which included: the biographical background of the interviewee (both academic record and current position), their motivation in using sonification, their vision of sonification, their specific design process and validation process, as illustrated through specific cases. The topics were handled in a free-form, closer to a narrative interview (Jovchelovich and Bauer 2000) than a structured interview, with some questions being more relevant than others in different phases of the interview process, according to the status of my research and the specific activity and expertise of the interviewee as in Supper cit., p.32. Additionally, the order of the questions was not set and often questions were not even explicitly formulated (Jovchelovich and Bauer cit., p.61), but rather used as a generic fil rouge to navigate the conversation.

Interview analysis: methodology- I approached the analysis of the qualitative material obtained through the interviews following classical text analysis (Bauer 2000) and qualitative text analysis (Kucartz 2014).

The process involves roughly three phases:

- Preliminary analysis based on pre-determined categories.
- Deep reading analysis, identification of emerging sub-categories.
- Identification of dimensions i.e., attributes to the sub-categories that define the range along which a category varies (Strauss and Corbin 1998, p.101).

Figure 17 recaps the qualitative text analysis iterative process as it appears in Kucartz cit., p.40.

The interviews, which took place remotely via Skype or Zoom, were first transcribed and carefully read for a general understanding. In this phase, a summary of each interview, which contained both emerging topics and a list of emerging keywords, was prepared. A second, close reading process, which aimed to interpret the text and identify emerging topics, followed. The text analysis used a mixture of deductive and inductive methods (Kucartz cit., Chapter 3). The deductive method was used to group the text around four previously defined categories. These four categories were constructed based on the existing guiding hypothesis of my research and in particular: around the three above-mentioned areas of investigation (sound design practices, design process and validation process) and the intentionality criteria.

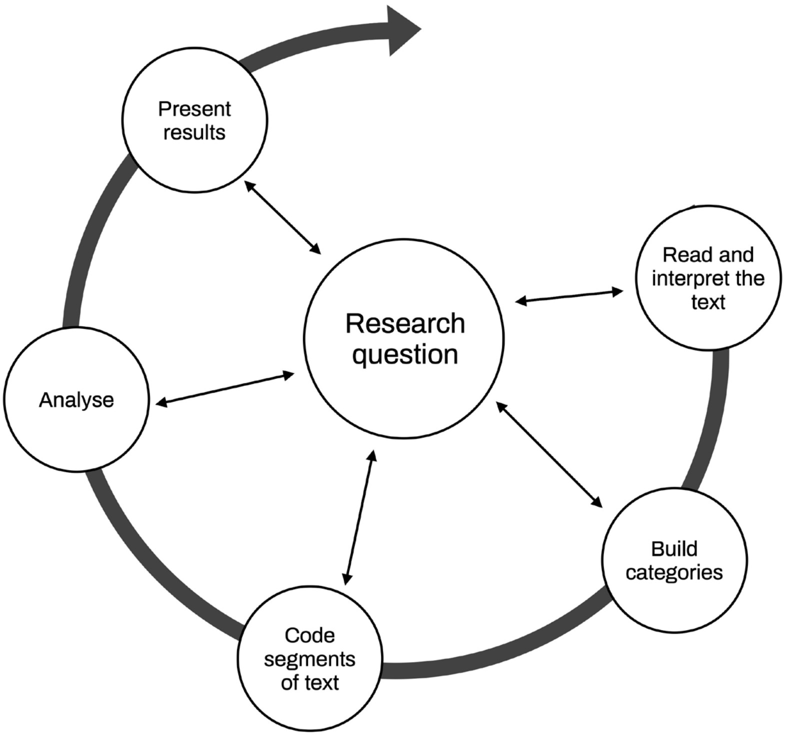


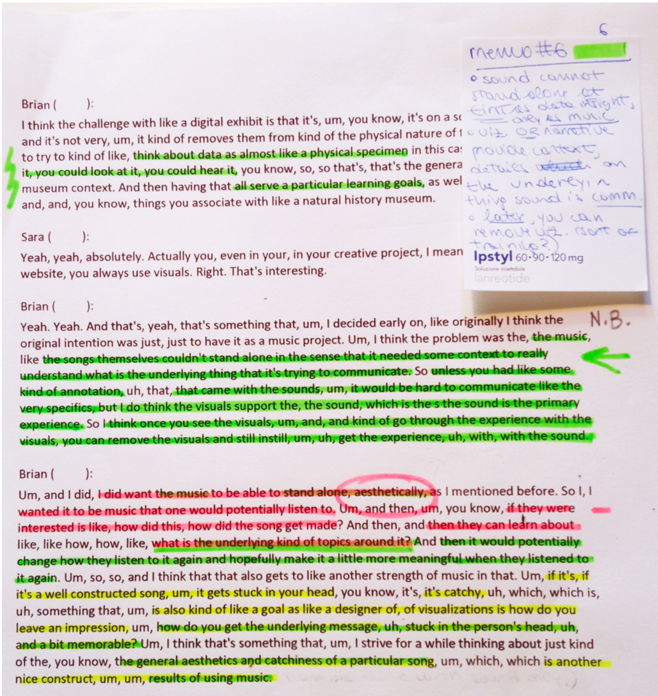
Fig. 17: The qualitative text analysis process, Kucartz 2014.

A fifth category was used to explore the 'theoretical framework' i.e., the conceptual space in which each author moved in his or her approach to overarching questions such as the value of music, theories of knowledge and cognition, the distinction between art and science. The four deductive categories used for text analysis were:

- Sound design.
- Design process (which also included the validation process).
- Intentionality.
- Theoretical framework.

From these four pre-determined categories, new sub-categories were extracted through an inductive process. Sub-categories were later defined through attributes (dimensions, in the definition of Bauer, cit.) that I later called 'decisions'. The following paragraphs describe this process in depth.

Pre-determined categories- On a printed version of the transcribed interviews, a series of a memos were taken during the close reading process and each was assigned to one of the four pre-determined categories. Memos were assigned a color - that was different for each category - in order to facilitate clustering at a later stage. Figure 18 shows a sample of this process with a memo corresponding to the categories of sound design (in yellow), sonification (in green) and approach to sound and music (in pink).


Brian ( ):

I think the challenge with like a digital exhibit is that it's, um, you know, it's on a sc
and it's not very, um, it kind of removes them from kind of the physical nature of t
to try to kind of like, think about data as almost like a physical specimen in this ca:
it, you could look at it, you could hear it, you know, so, so that's, that's the genera
museum context. And then having that all serve a particular learning goals, as wel
and, and, you know, things you associate with like a natural history museum.

Sara ( ):

Yeah, yeah, absolutely. Actually you, even in your, in your creative project, I mean
website, you always use visuals. Right. That's interesting.

Brian ( ):

Yeah. Yeah. And that's, yeah, that's something that, um, I decided early on, like originally I think the original intention was just, just to have it as a music project. Um, I think the problem was the, the music, like the songs themselves couldn't stand alone in the sense that it needed some context to really understand what is the underlying thing that it's trying to communicate. So unless you had like some kind of annotation, uh, that, that came with the sounds, um, it would be hard to communicate like the very specifics, but I do think the visuals support the, the sound, which is the s the sound is the primary experience. So I think once you see the visuals, um, and, and kind of go through the experience with the visuals, you can remove the visuals and still instill, um, uh, get the experience, uh, with, with the sound.

Brian ( ):

Um, and I did, I did want the music to be able to stand alone, aesthetically, as I mentioned before. So I, I wanted it to be music that one would potentially listen to. Um, and then, um, you know, if they were interested is like, how did this, how did the song get made? And then, and then they can learn about like, like how, how, like, what is the underlying kind of topics around it? And then it would potentially change how they listen to it again and hopefully make it a little more meaningful when they listened to it again. Um, so, so, and I think that that also gets to like another strength of music in that. Um, if it's, if it's a well constructed song, um, it gets stuck in your head, you know, it's, it's catchy, uh, which, which is, uh, something that, um, is also kind of like a goal as like a designer of, of visualizations is how do you leave an impression, um, how do you get the underlying message, uh, stuck in the person's head, uh, and a bit memorable? Um, I think that's something that, um, I strive for a while thinking about just kind of the, you know, the general aesthetics and catchiness of a particular song, um, which, which is another nice construct, um, um, results of using music.

Fig. 18: Example of qualitative text analysis based on pre-determined categories.

Memos were later moved to a mind map (see Figure 19) which helped identify, in an iterative process, a series of emerging sub-categories within each of the four pre-determined categories. Memos assigned to the same category during the close reading phase were grouped together and informed the definition of a series of sub-categories, each of which were connected to one of the main categories as shown in Figure 19. In a feedback loop which required several rounds of close reading of the text, memos were moved from one sub-category to another in order to start structuring a sonification design space.

In Figure 19 we can see the emerging structure of the design space. The four main categories branch out into sub-categories which were obtained through the analysis of firsthand comments from the seven interviewees. Some of these comments – the most relevant – are reported in the mind map for the clarity and transparency of the inductive process and will be illustrated in more detail in the next sections.

Emerging sub-categories- As shown in Figure 19, four sub-categories were identified within the theoretical framework: cultural divide; role of aesthetics; musical sonification; embodied sonification. In the intentionality category, interviewees shared their vision on the main goal of producing sonification: as a tool for investigation; as a tool for creative production; as a tool to increase audience engagement. Five sub-categories emerged within the design process: definition of the use case; prototyping and iterating; relationship with the phenomenon; explainability; multi-modality. Three sub-categories are associated with sound design: mapping strategies; listening experience; sound and music assets.

Dimensions- Kuzcak (cit.) defines dimensions as attributes of a category (or sub-category) which specify the range along which these categories vary. At some point during the analysis process, it was clear that sub-categories would not be sufficient to define the extent of the approach to sonification that was emerging from the interviews. Another layer of explanation was needed. At the same time, dimensions did not seem to be the appropriate tool. In fact, the identified sub-categories did not need to be placed on a gradient scale, which dimensions require.

On the other hand, a specific set of concepts seemed to emerge from the sub-categories 'design process' and 'sound design'. These concepts identified actions – or rather, decisions – that authors take during the process of creating a sonification. Figure 19 shows these dimensions as they emerge within the 'design process' category and the 'sound design' category. In particular, the sub-category 'mapping strategies' is further defined by decisions on what we called the sonification discourse; the type of sounds used in the sonification; the behavior of these sounds; the function of sounds within the mapping strategy. The sub-category 'multimodality' is defined by how sonification can combine with other sensory modes of experience, particularly visualization and interactivity. I will call these emerging concepts 'decisions', as they primarily refer to the intentional choices of the sonification designer during the design process.

Fig. 19: Interviews' analysis mind map.


I'm interested in sonification for human rights investigation (easier to record than to film)
It's a way to engage with a phenomenon and do something interesting for people to experience
as a tool to experience data in new ways
Sound represents the overall experience and feeling. It should not try to be mapped to hard values
We need to start with something that sounds interesting and pleasant then manipulate some parameters to provide information on the data, but it still sounds nice.
It's hard to package actual data (numerical) into sound: 'you get a sense of something but not analytical values
Structure/ "sonification discourse"
I like to represent the underlying data source in the sound and the data itself as a kind of fluctuation of the sound
Data behaves in a certain way and this behaviour should be reflected in the behaviour of the sound that comes out
Literal sounds can be too neutral
I try to find mapping strategies that are generalizable
Sound behaviour
Type of sounds
Choice of sounds is a preliminary creative constrain, a decision that has to be made explicitely
sound is a relatively inefficient way of delivering a lot of information especially with concrete numbers
Sometimes the result of a sonification is very complex acoustically and it's difficult for the listener to distinguish changes that are in the sound vs changes provoked by the dataset
Issues
1. Mapping strategy
Often, arbitrary mapping to musical frameworks takes a lot for granted. How a musical structure works is not reflected in how the data works.
Sounds are aesthetically related to the data and the underlying topic
Function of sound
The mapping problem is the old mind-body problem. There are objective data and subjective interpretation of sound, and how are you supposed to bridge the two? You have to change the way you look at it.
Data are not neutral so sounds chosen to represent it should not be neutral either.
You can use metaphors to take a comforting sound (like sea sound) and fill it with a different emotion
Design Space
SOUND DESIGN
Use of 'neutral' sounds might misrepresent the complexity of the phenomenon.
The main issue is what sounds do you use? what is the meaning and what do you want people to feel? And that is really, really difficult.
You can take a familiar comforting sound and fill it with a different emotion.
The context of experiencing the sonification impacts the design (HP, physical space, web space etc)
1. Definition of the use- case
Sonification works well if you have a specific constrain to use sound
Film music is a good example of how to communicate in a good way with sound
3. Sound and music assets
You are supposed to listen to the sound and understand specific data. It does not really make a lo of sense often.
2. Listening experience
I practice 'reduced listening' myself, to evaluate whether the message I want to communicate arrives to my listener
there is a lot of associations going on if people listen to concrete sounds by recognising the sounds or its properties
Sometimes sound increases performance in association with visuals, you cope better with the amount of information
We cam easily detect changes in sound patterns in an immediate, visceral way but usually for very simple datasets
We have to ask ourselves how should the listener be feeling about this specific dataset?
How does the listener know what the set of rules is (to make it sound musical) and where the data fit in?
Music has the unique capacity to change your mood. The listener will feel something.
Sound can contribute to give the time perspective that you lose in static visualization.

Results. Towards the definition of a sonification design space-
Figure 20 proposes a schematic view of the four categories with their respective sub-categories and associated decisions. In a clockwise order, the reader can navigate the design space from the abstraction of the theoretical framework, the backdrop that inspires the daily practice of the designer, to the practicalities of the individual steps and specific decisions that have to be taken 'to get the project done'. These steps and decisions are the result of a mainly tacit knowledge formed through design practice. The interview process aimed to reveal this knowledge, thus making explicit the rationale behind the author's approach to sonification. The extent to which this rationale can be generalizable and embedded in a tool to support designers in the production of a sonification will be explored in the following sections.

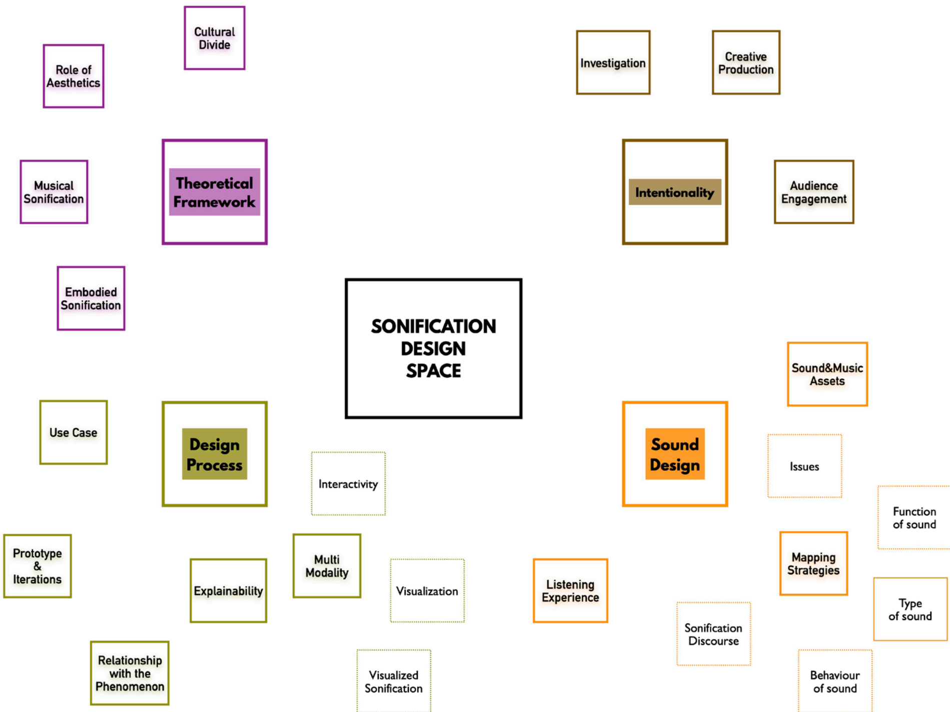


Fig. 20: The sonification design space.

I will first describe in more details each quadrant, in clockwise order, before presenting a brief summary of the results.

Theoretical framework- When reflecting on his approach to sonification, one of the interviewees stated, not without a certain bitterness, that "You tend to find that old divide [...], there is a definite divide between an engineering and sciencey approach and an arts approach. There is a divide between art people and technology people, and they don't cross over." He is referring to several frustrating situations ("It was really tough actually") in which "no one in the art group would do anything that had a "technical slant" whereas "Technical people would be like, well, all our applications are purely scientific or else they are commercial". The interviewee, an interaction design researcher, concludes that perhaps "Design should unite the two and provide this middle ground". The idea of a *cultural divide* within the sonification community, between an approach that has hard sciences as a model versus an approach that stems out of art, has been extensively investigated by Supper (cit.). I decided to include it as a sub-category as it reflects a polarization that can affect the emergence of sonification as an independent discipline built on a common ground.

The *role of aesthetics* seems to involve another traditional divide that radicalizes opinions and insulates people within the two separate poles of arts, on the one side, and science on the other[18]: the (often implicit) definition of aesthetics as being related to beauty (and beautification) or the definition of aesthetics as the act of knowing through the senses. The interviewees referred to the role of aesthetics in sonification in both ways. On one hand, some of them claimed that "We must have an aesthetic approach to sonification, you need to be able to say, this sounds nice" and that "Sonification should be pleasant to listen to", thus explicitly referring to the aesthetic experience as something "pleasant, interesting and entertaining" that makes it possible for the "informative part to work". On the other hand, some authors insisted during the interview that "There is a lot more to aesthetics than just sounding good. It's about the choices that the designer makes and how these influence your understanding of data.", even claiming that "What many people mean by aesthetic is 'cosmetic'".

18.
For an enlightening and entertaining narrative of the divide between science and humanities, the reader can refer to C. P. Snow, 'The two cultures' (Snow 1959).

Embodied sonification reflects a specific approach to the use of sound to represent data which stems out of the philosophical debate on a cognitivist approach versus an embodied approach to the theories of knowledge. Embodied sonification is a framework that leverages the existing conceptual metaphors through which we make sense of reality (Lakoff and Johnson 1981) to design sonifications that, as one interviewee described them, “use sound as a connective tissue between the data and the listener’s previous experience of the world.” Embracing a general framework such as embodied cognition also provides a backdrop for the sonification design process: “I take the data on one side and then I explore models of how a person might make sense of them in terms of embodied cognition.”

One of the authors (an interaction designer with a background in media studies and self-taught music production) defined what he does as *musical sonification*, an approach that helps him use “shortcuts to reach a user or a listener” that music, a discipline that creates emotional experiences and has “the unique capacity to change your mood” can provide. “If we had a musical approach to sonification, we would end up doing nicer or better, more useful sonifications”, he claims.
On the other hand, another interviewee, a composer and professor of composition, pointed to the issues we have when defining what music is: “If by music you mean conventional tonal harmony, are you limiting it all to a major scale? That seems reductive. Also, there’s huge differences across musical cultures in the world”. Furthermore, using musical dimensions to map data to sound could complicate rather than simplify the experience for the listener, as music composition has its own rules and associations that would not be reflected in the data”. With the consequence that “Often, arbitrary mapping to musical frameworks takes a lot for granted. The way a musical structure works is not reflected in how the data works”.

Intentionality- In this analysis, we considered the criterion of intentionality (that we defined in the first section of this chapter) as an indicator of the overarching motivation of the interviewee in using sound to represent data. Considerations on the reasons for using sonification emerged freely during the interviews. Reflections were grouped around three main clusters:

- Sonification to engage the audience, especially on socially relevant issues.
- Sonification as a personal tool for creative productions.
- Sonification as a tool for investigating a phenomenon in order to gather and communicate information.

Some comments stand out more than others. One interviewee, a mathematician and self-taught musician, claimed he uses sonification “to reconcile with numbers”. Another uses sonification “to do algorithmic composition, with a message”. But it is in the sub-category of sonification as a tool to engage the audience that the role and the responsibility of the designer emerge: statements such as “I want to ‘disturb’ the listener in a way that makes them obliged to engage with the phenomenon” or “Sound should ‘break the day’ of the audience, capture their attention on a specific phenomenon, make it memorable” are a testimony to the potentiality of sound “To emotionally engage with a phenomenon and with numbers, especially in socially relevant issues”.

Sound Design- To recap, the sub-categories of sound design are mapping strategies, listening experience and sound and music assets. As these sub-categories are defined by different decisional areas (‘decisions’), a separate paragraph will be dedicated to each of them.

Mapping Strategies

As shown in Figure 20, I identified four main decisional areas, or decisions under the sub-category ‘mapping strategies’: sonification discourse, sound types, sound functions, sound behavior. A somewhat separate area groups together statements that raise issues on the overall process of mapping data to sound.

I grouped under ‘sonification discourse’ design decisions that pertain to the articulation of the message contained in a sonification. These decisions answer questions such as: how do we organize the sound material in order to serve the specific purpose of our sonification? How does the message we want to communicate to the listener influence the structure of the sonification itself?

In Simon Emmersons’ seminal ‘The Language of Electroacoustic Music’ (Emmerson 1986), it is the advent of tape music (music in which sounds are not produced live, but are instead recorded, manipulated and played back through a physical support like a tape, a CD, a computer) that expanded the possibilities of how a composer shapes the relationships “of the sounds to associated or evoked images in the mind of the listener”. Recording technologies have liberated the composer from the constraints of the causal relationship between a sound and the gesture that produces it. In traditional music, the public could see a pianist touching the keys at the same time as they heard the sound – a direct causal relationship that limited the possibilities of associating the sound they heard to a different cause. When a sound is recorded, manipulated and subsequently played back through speakers, the decision (and the responsibility) on how much the real sound source should be disclosed to the listener lies in the hands of the composer. Emmerson calls the area pertaining to decisions on how we communicate the message to the audience and which 'mental images' we want to evoke 'discourse'. These decisions do not specifically refer to the choice of sound material (what we will later call the 'sound types') or to choices around the compositional organization of such material (what we will later describe as the 'sound functions' and the 'sound behavior'). Nor do they relate to specific listening attitudes (what we will call the 'listening experiences') that an author seeks to elicit in her public. Rather, they relate to the overall approach to the communicative relationship that a composer establishes with the audience through sound, and that reflect the intention of the composer herself in shaping this relationship. As one of our interviewees described it, the author of a sonification can "Start with something that sounds interesting and pleasant then manipulate some parameters to provide information on the data" or she can use sonification "to represent the overall experience and feeling of a phenomenon", without "trying to map it to hard values". In other words, we can use sonification to tell a story or we can use it to communicate specific analytical values.

Decisions on the 'sound types' reflect the different approaches of interviewees to the choice of the raw sonic material they use to build their sonifications. There is not general agreement, to my knowledge, on how to define this raw material when, as we described in the previous paragraph, it is abstracted form the ephemerality of its natural occurrence and stored on a technological support for its transformation into a new sonic artefact. The interviewees referred to their working material as 'sound content', 'content', 'sound material', or simply as 'the sounds'. Traditionally, classical electroacoustic music followed Schaeffer in identifying the 'sound object' (the *Objet Sonore*, Schaeffer cit.) as the basic unit of electroacoustic music composition. Technically, sound engineers and music producers talk about 'samples' or even simply 'sound files' when they describe the raw material used in their daily work. I labelled as 'type of sounds' the area in which our interviewees describe their decisions around which sounds to use in a sonification, a “preliminary creative constraint that has to be made explicitly”.

Sound is a time-based phenomenon, in that, unlike a visual object, a sound object can only occur in time. When approaching a sonification, authors have to take decisions on how the type of sound they choose will exist in time. I called the area to which these decisions relate ‘sound behavior’. In a sonification, the rules on the changes of sound over time acquire an increased importance as, typically, phenomena that data represent do change over time. As one of our interviewees put it: “Data behaves in a certain way and this behavior should be reflected in the behavior of the sound that comes out”. What are the options of the designer when it comes to shaping the temporal relationship between data and sound? Strategies can be manifold and decisions must be taken during the design process. At two opposite extremes, one of our authors related that he likes “to represent the underlying data source in the sound and the data itself as a kind of fluctuation of the sound” while other try to “to find mapping strategies that are generalizable”.

“The main issue is what sounds do you use? What is the meaning and what do you want people to feel? And that is really, really difficult.” declares one of my interviewees, a designer and activist who works with data collection and data visualization (occasionally sonification) to investigate abuses of power in the context of human rights violations. We call this decisional area ‘sound function’. Here the designer has to answer the question: what is the function of the specific sound I choose in relation to the specificities of the phenomenon I am representing? For one of the authors I interviewed, “Sounds are aesthetically related to the data and the underlying topic” whereas another author highlights the risks of relying on musical sounds and structures in which “Arbitrary mapping [to musical frameworks] takes a lot for granted. The way a musical structure works is not reflected in how the data works”. Particular attention is given to the role of metaphors (“You can use metaphors to take a comforting sound (like the sound of the sea) and fill it with a different emotion”) and to the hidden ambiguity in trying to maintain sound as neutral (“Use of ‘neutral’ sounds might misrepresent the complexity of the phenomenon”, “Data are not neutral so sounds chosen to represent them should not be neutral either”). In the next chapter, I will describe how the functions of sound are interpreted in semiotic terms and used as a building block for the sonification design canvas.

As mentioned, I decided to keep in a separate group comments which highlighted issues in the mapping process - issues that are not new to the sonification community (see p.32). Not all our interviewees actively belong to the sonification community per se, as sonification (or sound design) is not their main occupation. Still, they share similar concerns when they say that “Sometimes the result of a sonification is very complex acoustically and it’s difficult for the listener to distinguish changes that are in the sound vs changes provoked by the dataset”. One author, who in his profession mainly uses visualization as a tool for data representation, claims that “Sound is a relatively inefficient way of delivering a lot of information especially with concrete numbers” while on the other hand it has the unique capacity of “making you feel something”. After all, “The mapping problem is the old mind-body problem.” - says another interviewee - “There are objective data and subjective interpretations of sound, and how are you supposed to bridge the two? You have to change the way you look at it”. Could design provide the necessary tools to pivot a theoretical conundrum into a design problem where explicit decisions are taken to solve the specific needs of a specific audience in a specific context?

Listening Experience

On various occasions, interviewees highlighted the role of the listener and her appreciation of the relationship between sound and data in a sonification. Imagining the listening experience of your audience, according to one of the interviewees, helps answer the question: “How should the listener be feeling about this specific dataset?”. Wirfs-Brock et al. (cit.) describe how they used self-reflection techniques during the sonification design process in order to listen “through the ears of an imaginary audience”. Part of the design process they propose consists in making explicit the implicit mental models of the audience that a designer conjures up, similarly to what happens in HCI with the definition of personas in user-centered design process.

During the interviews, the role of one’s proxy emerged, at times surrounded by discouragement and skepticism (“You are supposed to listen to the sound and understand specific data. It does not really make a lot of sense often.", "How does the listener know what the set of rules is to make it sound musical and to establish where the data fit in?"), other times as a specific design strategy that helps shape the sonification process ("I practice 'reduced listening' myself, to evaluate whether the message I want to communicate reaches my listener").

The reference to 'reduced listening' (a close-listening practice mainly performed by professional sound designers and composers when they focus on the acoustic properties of their work to evaluate its quality), together with references to the causal dimension of acoustic phenomena ("There are a lot of associations going on if people listen to concrete sounds by recognizing the source or its properties"), open the door to introducing listening modes as they are defined in literature by Chion (1994 cit.) and Gaver (1993). In the following chapter we will describe how these modes are used to build the 'listening experience' block in the sonification canvas.

Sound and Music Assets

Literature on sonification has put some effort into identifying characteristics of sound that would give it specific advantages in the representation of data. In sonification literature, the most quoted characteristics are the capability of the human ear to perceive changes in sound events and simultaneously distinguish between multiple layers in the same acoustic phenomenon such as timbre, pitch, amplitude, rhythmical patterns (to name a few, Scaletti and Carig cit.; Barrass 1997 and 2003, cit., Vickers et al. cit.; Lenzi et al. cit.). In practice, it is still difficult to distinguish scientifically proven effects sound has on human perception from orally transmitted knowledge in sound design and music production– knowledge that is often tacit and closer to the tradition of fine arts and craftsmanship.

What our authors are referring to when talking about the added value of sound and music seems to be a mixture of self-reflection on what 'works' in their daily practice with sound and the results of basic research in sonification. "We can easily detect changes in sound patterns in an immediate, visceral way" says one interviewee, repeating one of the mantras of sonification literature. "But it usually works only for very simple datasets." he adds, in what seems to be a comment based on first-hand experience. "Film music is a good example of how to communicate in a good way with sound" says another interviewee, echoed by another who still claims "Music has the unique capacity to change your mood. The listener will feel something". An interviewee who is directly involved in basic sonification research i.e., research that aims at investigating and validating specific characteristics of sound that can be later integrated in 'applied' sonification projects, reports that "Sometimes sound increases performance in association with visuals, you cope better with the amount of information".

Design Process- To recap, the sub-categories for design process, as illustrated in Figure 20 are: definition of the use case; prototyping and iterating; relationship with the phenomenon; explainability; multi-modality.

Definition of the use case

"Sonification works well if you have a specific constraint to use sound", says one of the interviewees, while another indicates that the context in which the audience experiences a sonification (whether on the web or in a physical space, through headphones or through speakers) "impacts the design" of the sonification itself. Using a term that is common in design practice, we grouped such statements under the area of the use case definition in which a designer defines users, goals, contexts and constraints of her design action.

Prototyping and Iterating

Most of the interviewees seem to engage – explicitly or otherwise – with a typical design process (such as those described in Chapter 2) when they describe how they move from concept to artifact. Not without struggles ("It is very difficult to work with real people"), some of them "explore from very early stages how a sonification works with real people and preferably in real world situations". Sometimes it takes a leap of faith to borrow practices from other fields, such as interaction design: "One of the interaction design guys said, why don't we do an example? Let's see how it works. What do you mean? Should I just do it? And this was a bit scary." Interestingly, some of the authors go through an early exploratory stage of prototyping that either uses sound differently ("I usually start with data driven music that never sees the light of the day to explore dimensions I could map the sound to.") or uses visualization in order to familiarize with the dataset and gather ideas on how to approach the sonification ("I visualize data first to better understand what I want to communicate with sound."). This second approach is echoed by other sonification authors such as Wirfs-Brock and colleagues (cit.). Validating how generalizable these sub-categories and their respective decisional areas are will be one of the long-term goals of the sonification canvas we propose in Chapter 3.

Under the sub-category *Relationship with the phenomenon* I grouped comments that relate to the approach authors take when they tackle the connection between sound, data and the real-world phenomenon. These three elements can be approached in a different order. One author says "I start with a dataset then go back to the underlying topic and choose what I want to communicate about it and how". Another first chooses a topic "I want to know more about then find a dataset that is representative of it.". Other authors, as illustrated earlier in this chapter frame the relationship within a specific theoretical approach: "I take the data on one side and then I explore models of how a person might make sense of it, in terms of embodied cognition". One of the interviewees highlights issues of neutrality when it comes to design decisions: "Sonification cannot be neutral as we have to take at least two decisions: which data we sonify and with which sound. Choosing randomness would still be a choice".

The sub-category I called *Explainability*, referencing a term more commonly used in computer science[19], (Doshi-Velez and Kim 2017) refers to a still rather underrepresented topic ("Sonification practitioners have focused more on the generation of sounds from datasets than on teaching people how to listen to and interpret them", say Wirfs-Block and colleagues, cit.) which emerged during the interview process. This topic addresses questions such as: does a sonification have to be explained before the audience experiences it? And if so, how much, with what means and with what expectations? According to one of the authors "Sonification can stand alone only if previously experienced with annotations or visuals" while another one claims that "transparency and documentation on the decisions taken during the process is key to make a sonification more robust".

Multimodality is developed in three decisional areas: 'visualization', 'interactivity' and 'visualized sonification'. "I don't use sonification by itself but always with some other experience: you need to think about

19.
The term 'explainability' usually relates to the concept of 'explainable Artificial Intelligence' i.e., the collection of methods and techniques in the application of artificial intelligence technology (AI) such that the results of the solution can be understood by humans. It contrasts with the concept of the "black box" in machine learning where even its designers cannot explain why an AI arrived at a specific decision. From Wikipedia, the free encyclopedia.

interaction to keep the interest of the public" says one of the authors, a sonification practitioner and researcher who works mainly in collaboration with interaction designers. The relationship with visualization is somewhat more complex, with the visual support that can be used both as a backdrop for the sonification or as the main means of representation. Interviewees seem to use it both ways, with one author noting that "Visualization (or annotations) can provide the context and the narrative to a sonification" while another states that "Sound can add pleasure and fun to reinforce the visualization of hard values". With 'visualized sonification' I identified the circumstances in which a sonification project cannot stand alone due to the specific nature of the medium, for example, on a web page: "Web communication needs a visual/interactive experience to keep the audience's attention".

These emerging dimensions – areas of the design space where the interviewees, i.e., authors of sonifications, take conscious decisions - were used as a basis to define the building blocks of a design tool – the sonification canvas – that will be described in detail in the next chapter.

CHAPTER 3 —
Primo tema.
The data sonification canvas.

Soundtrack: *T. Monk,Thelonious himself.*

Building the blocks

At the end of the qualitative analysis of the expert interviews I identified several areas –for both the sub-categories and the decisions level - in which authors seem to take conscious decisions that determine the design of the sonification. Specifically, these areas are 'definition of the use case'; 'relationship with the phenomenon'; 'listening experience'; 'multi-modality' and 'mapping strategies'. The latter is further developed in 'type of sounds', 'function', 'behavior'.

I used these seven decisional areas to draft the building blocks of a design tool, the sonification canvas, which aims to support designers of various disciplines (communication design, information design, service design, etc.) in using sonification as a method to represent and communicate data to an audience. In Figure 21 the reader can see an early prototype of the canvas. The prototype was later refined through a three-hour workshop in which five communication designers (with no specific expertise in data sonification) were asked to catalogue existing sonification cases using the canvas' decisional blocks. The exercise aimed to assess the coherence of the blocks, the definitions and examples used to explain them, and the overall potentialities of the canvas in accounting for the characteristics of a sonification project.

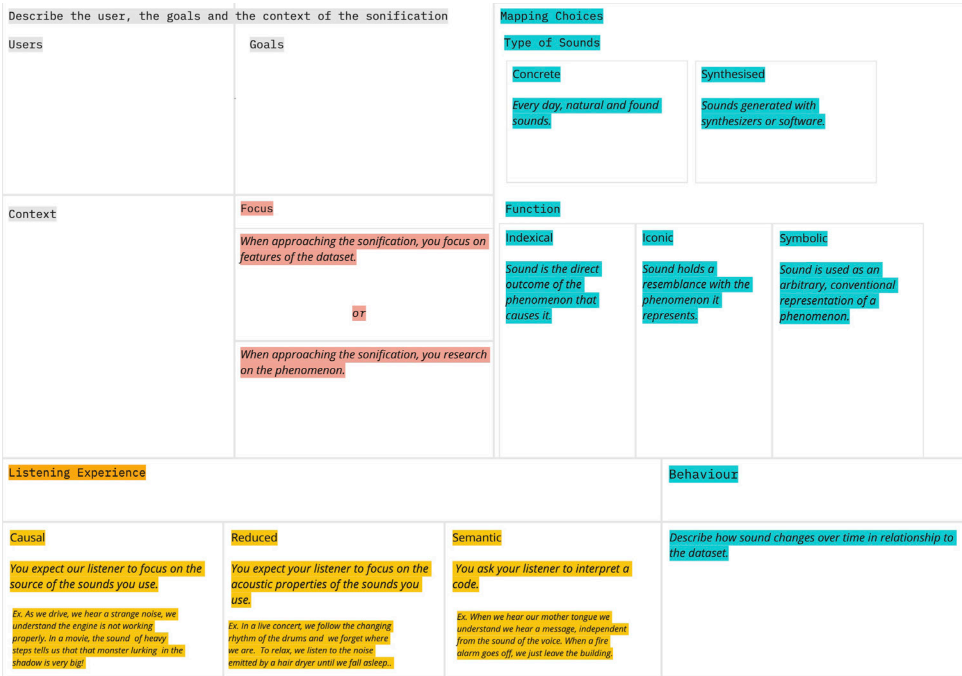


Fig. 21: Data sonification canvas, first prototype.

In the following paragraphs, I will illustrate each block before describing in detail the workshop. Feedback from the workshop helped me clarify the language (both the description and the examples) of some of the blocks. New blocks were added to reflect other design aspects that emerged during the workshop. It did not, however, involve a radical review of the canvas itself, which maintained the same structure and, in general, the same blocks as the first version.

Block 1. Describe the user, the goals and the context of the sonification- This is the first block, with which I imagine a designer will start her journey in data sonification. The block requires defining the design problem, describing in detail the user of the sonification, the specific constraints of the project, the context in which the sonification will be experienced and finally, it requires making explicit the goals the designer wants to achieve with the sonification.

Block 2. Mapping choices- Following the canvas in clockwise order we encounter one of the most important blocks for the author of a sonification, where all the practical decisions that will determine the strategy for translating data into sound are taken. Here I gathered the decisional areas that are grouped under the sub-category “sound design”.

Type of Sounds

This block refers to the choice of sound material i.e., the actual sounds that the designer will source to create the sonification. This choice will depend on a variety of factors, from design constraints to specific requests from the audience, to the personal aesthetic and expressive preferences of the designers. I chose the terms synthesized and concrete to represent two opposite poles in the nature of the sound material that the designer can choose. We use synthesized sound when the material is artificially created using specific electric or electronic devices. In early electroacoustic music, synthesizers were complex machines based on electric oscillators that could generate waves within the frequency range of human hearing'[20].

20. A good resource on the history of synthesizers is https://making-music.com/quick-guides/oscillators/.

These waves would then be added or subtracted to/from one another, then filtered or reverberated to obtain complex, artificial sounds. Several synthesis methods have been developed during the course of the 20th Century, particularly since the 1980s with the advent of digital synthesis and the resulting increase in the complexity and quality of the sounds produced. Today, commercial and professional audio software offer an incredibly sophisticated range of sound synthesis tools. By manipulating a high number of different parameters, the designer can create her own sounds or use high quality factory libraries. At the opposite pole, concrete sound refers to material that is sourced from the very world we live in. Sounds of natural origin (such as those sourced from animals or the environment) and sounds of anthropic origin (such as those obtained by manipulating everyday objects or from human voices, machines musical instruments) were recorded first on analog (the vinyl), then digital (the CD and the hard drive) support. When abstracted from their natural environment, they become the building blocks the sound designer or the composer can process, edit and assemble in his creations.

Function of Sound

In the previous chapter, decisions on the function of sounds referred to comments about the relationship of sound with the phenomena it represents. In the canvas, this relationship is illustrated in semiotic terms: sounds can be used as indexes, icons or symbols of the real-world element they represent. Over the years the field of auditory display has developed a specific categorization of sound functions in semiotic terms, a categorization that is usually leveraged to categorize and improve the communicability of auditory displays (Jeon 2015). Referencing the theoretical framework of Pierce's semiotics (1902) and the categorization of signs as icons, symbols and indexes, sounds used in the design of auditory displays are defined respectively as 'auditory icons', 'earcons' or, more recently, 'spearcons', 'lysicons' and 'spindexes' (Jeon cit.) depending on their indexical, iconic or symbolic function.

Auditory Icons

Auditory icons were first defined by Gaver at the end of the 1980s (specifically in 1986 and 1989) as "caricatures of naturally occurring sounds that could be used to provide information about sources of data." (1989 cit., p.167). Gaver proposes the use of auditory icons in order to "emphasize the role of sound in conveying information about the world to the listener". Through auditory icons, listeners gain information on the real-world source that produced the sound. In designing an auditory icon, we map characteristics of a real-world phenomenon to sound. One advantage of using iconic sounds, when compared to the use of arbitrary conventions, is that "if a good mapping between a source of sound and a source of data can be found, the meaning of an auditory icon should be easily learned and remembered." (1989 cit., p.169).

In the design process, "it should be possible to create auditory icons that represent the objects and actions of the computer world in an intuitive way, simply by mapping them to the objects and interactions of everyday sound- producing events. The appropriate mappings should be obvious" and consequently "the relations between them and the information they are to convey should be obvious." (1986 cit., p.71). Gaver, who was at the time also working as a sound designer and researcher at Apple Computers, gives the following example: "The file hits the mailbox, causing it to emit a characteristic sound. Because it is a large message, it makes a rather weighty sound. The crackle of paper indicates a text file- if it had been a compiled program, it would have clanged like metal. The sound comes from the left and is muffled: The mailbox must be in the window behind the one that is currently on the left side of the screen. And the echoes sound like a large empty room, so the load on the system must be fairly low." (1989 cit., p.166-167). Interestingly, this usage of sound mirrors what Chion describes as ‘materializing sound indices (m.s.i.)’: “The materializing indices are the sound’s details that cause us to “feel” the material conditions of the sound source and refer to the concrete process of the sound’s production. They can give us information about the substance causing the sound—wood, metal, paper, cloth—as well as the way the sound is produced—by friction, impact, uneven oscillations, periodic movement back and forth, and so on.” (1994 cit., p.114). Critics to Gaver’s definition of auditory icons argue that he “obviously confuses index and icon when he claims that ‘iconic mappings are based on physical causation’ and ‘its characteristics are causally related to the things it represents’. This is true for indices, but not for icons, which are not based on causality but on similarity.” (Oswald 2012, p.39). Petocz and colleagues (2008) proposed renaming Gaver’s auditory icons ‘auditory indices’. Despite these contradictions, ‘auditory icon’ is still a generically accepted term in the data sonification community and refers to the design of sounds that share some of the characteristics of the phenomenon they represent, as opposed to sounds that rely on an arbitrary relationship, such as earcons.

Earcons

The term ‘earcon’ was first coined around the same time as Gaver’s definition by Blattern, Sumikava and Greenberg (1989) who defined earcons as “auditory signs based on musical principles – short micro-compositions of only a few notes length.” (Oswald cit., p.40). McGookin and Brewster (2011, p.339) expands on the definition stating that earcons “can be thought of as short, structured musical messages, where different musical properties of sound are associated with different parameters of the data being communicated. The key difference between these and Auditory Icons is that there is no assumption of an existing relationship between the sound and the information that it represents. This relationship must, at least initially, be learned.” From Gaver’s perspective, this is a disadvantage due to the arbitrary nature of earcons, which have a symbolic and not causal relationship with the real-world and as such cannot be understood intuitively. On the other hand, Oswald points to how some characteristics of music can be directly understood, across cultural differences and independently from training, such as the sense of tempo and the change of mood. Consequently, there is “a plethora of associations evoked by musical universals that can be utilized in earcon design in order to serve a communicative goal" (Oswald cit., p.40).

Spearcons, Lyricons and Spindexes

Spearcons (Walker Nance Lindsay 2006), Lyricons (Jeon and Sun 2014) and Spindexes (Jeon and Walker 2009) are recently introduced categorizations in which speech is used in different ways as an auditory cue for the design of auditory displays. Specifically, spearcons are produced by compressing spoken phrases; lyricons combine speech with earcons; spindexes are "brief non-speech auditory cue based on the pronunciation of the first letter of each item" (Jeon and Walker cit.) that the display is representing. Due to their nature, these elements seem to pertain to the realm of symbolic representations. However, Jeon (cit.) claims that "this symbolic relationship between the speech and an object is automatized based on life-long learning. With short training, the spearcon, 'apple' can have an indexical relationship with an actual apple due to its trace of the actual apple even though users do not recognize the original speech.". As such, spearcons, lyricons and spindexes should be considered as indexical in Peirce's terms i.e., as a necessary consequence of the phenomenon that caused them.

In a discussion of sonification and indexicality, we cannot fail to mention Vickers and Hogg's influential work (2006) on the aesthetic perspective of sonification (which they label 'ars informatica') in relation to sound art and music ('ars musica'). In this context, indexicality is defined as an overarching dimension in the continuum between ars informatica and ars musica, that defines "how strongly a sound sounds like the thing that made it." (ibid., p.231). In sonification, indexicality is related to the degree of direct data-to-sound mapping. A high indexicality is shown by sonifications in which the sound is derived directly from the data, whereas a low indexicality is shown in projects that use sound metaphorically.

As mentioned, the objective of the sonification canvas is to provide the communication designer with a tool to support her in the exploration of sonification as a means to represent data. After immersing myself in the complexities of an interpretation, in semiotic terms, of the use of sound for the display of information, I decided to take an approach that would be as generalizable as possible and as such, open to the design community rather than to the auditory display community. In the canvas, the possible functions that sound assumes in the relationship with the real-world phenomenon and data that it represents are defined as:

1. Indexical: sounds are indices i.e., a necessary consequence of the phenomenon that caused them. In a sonification, we use sound indexically when we ask the listener to detect the intensity of rain by listening to the sound that specific rain emits.

This definition echoes the definition of the sonification method called Audification (see p.17 for a definition). According to Kramer (cit., p.27) audification is "The direct playback of data samples". In the Sonification Handbook, Dombois and Eckert (cit.) expand on this definition stating that "In audification one can distinguish between different types of data that result in different types of sounds. Often, sound recordings themselves have already been named 'audification' if they have been shifted in pitch. Therefore, we want to include with our definition above all data sets that can be listened to, i.e., also all sound recordings themselves.". When, according to the canvas, a designer uses sounds indexically, we might imagine, in parallel with the concept of indexical design (Offenhuber and Telhan 2016), that she would mainly refer to recordings of the sounds emitted by the real phenomenon. In the example above, it refers to sound recordings of the rain, the intensity of which we want to analyze. This sound material is, to some extent, "autonomous and not under the control of the designer" (ibid., p.289). Although I do not wish to examine more deeply this complex matter given the goals of the present work, I suggest and hope that a real usage of the canvas and the collection of use cases where sound is used as an index could cast some light on the possible relationship between what we define here an indexical sonification, audification and more generally, indexical design.

2. Iconic: sounds are icons i.e., they are abstract representations of a phenomenon with which they hold a resemblance. In a sonification, we use iconic sounds when we represent machine to machine communication in a IoT network using sounds of human conversations.

3. Symbolic: sounds are symbols i.e., arbitrary representations whose meaning is conventional and has to be learnt by the listener. We use sound symbolically when we use a frog's call to

represent an incoming message on a mobile device.

Behavior

In the 'behavior' block in the canvas I expect the designer to describe how sound changes over time in relationship to the dataset i.e., the rules that govern how parameters in the data and parameters in the sounds are related.

Listening Experience- In the interview analysis, authors referred to a variety of experiences that a listener might activate when dealing with a sonification. From 'reduced' to 'causal' listening, different interviewees had different strategies to envision how their listeners "should be feeling about this specific dataset". This question is not uncommon in literature on auditory display and, to some extent, sonification. While developing the concept of auditory icons, Gaver states that "Perhaps the most important advantage of this strategy is that it is based on the way people listen to the world in their everyday lives." (1989 cit., p.168). The experience of everyday listening is described in contrast to musical listening, even if the two experiences do, in practice, coexist. "For instance, while listening to a string quartet we might be concerned with the patterns of sensation the sounds evoke (musical listening), or we might listen to the characteristics and identities of the instruments themselves (everyday listening). Conversely, while walking down a city street we are likely to listen to the sources of sounds – the size of an approaching car, how close it is and how quickly it is approaching – but occasionally we might listen to the world as we do music" (1993 cit., p.2). Chion (1994 cit.) dedicates a full chapter of 'The Audivision' to the three 'listening modes': causal, semantic and reduced.

Causal listening is the most common and "it consists of listening to a sound in order to gather information about its cause (or source)." (Chion ibid., p.25). Semantic listening relies on the interpretation of a code or a language to understand what we are hearing, as happens in spoken language or Morse code. Finally, reduced listening "focuses on the traits of the sound itself, independent of its cause and of its meaning. Reduced listening takes the sound— verbal, played on an instrument, noises, or whatever—as itself the object to be observed instead of as a vehicle for something else." (ibid., p.29). The same three listening modes are borrowed by Sonnenschein who also adds a fourth, ‘referential’ listening, which consists of “being aware of or affected by the context of the sound, linking not only to the source but principally to the emotional and dramatic meaning’ (cit., p.78). In the canvas, the designer is asked to describe the predominant listening attitude that she expects her audience to take on. From one side, I use Chion’s categorization of causal, semantic and reduced listening for its widespread recognition beyond the auditory display community and for a certain symmetry with the ‘function’s block (indexical, iconic and symbolic). At the same time, I decided to call the block, echoing Gaver, ‘listening experience’ (instead of listening modes) as it felt like a more appropriate term to address a non-specialized design audience.

Focus- Focus is the result of considerations that emerged from the interviews in areas such as ‘discourse’ and ‘relationship with the phenomenon’. It was my purpose to encourage the designer to reflect on the overall approach to the sonification at the very beginning of the project. In the interview analysis, these areas gathered comments on the possibility of communicating to the audience the general feeling of the phenomenon or, inversely, hard values from a dataset without referencing real-world events that generated the data. But how would I label this topic with a one-word definition? ‘Focus’, ‘approach’, ‘discourse’, were all on the table. Admittedly, this was the block for which definitions never seemed appropriate. For the first prototype of the canvas, I asked the designer to choose between ‘focusing on features of a dataset’ or ‘researching on the phenomenon’, hoping that feedback from real users during the workshop would help me clarify this.

The Canvas as a Tool for Designers

In order to validate the canvas prototype, I asked five members of the Density Design Lab at the Politecnico di Milano to take part in a two-hour activity in which the canvas was used to analyze 15 existing data sonification cases selected by me. Workshop participants were divided in two groups and introduced to the cases with the use of supporting material. Subsequently, they analyzed each case and filled in the canvas’ blocks i.e., they described the users of the sonification, the goals of the author, the context of experience, the focus, the types of sounds, their function and so on. The effort required the participants to take on the viewpoint of each project’s author and in doing so, describe and possibly explain the choices behind the design of the sonification.

Participants - communication designers mainly specialized in data visualization – had to interpret data sonification projects through the lenses of the canvas thereby using the canvas as an analysis (rather than as a production) tool. My hope was that through this exercise, inconsistency in the canvas' blocks would emerge and feedback would be generated on clarity, usefulness and on any missing elements. The choice of engaging information and data visualization designers, instead of data sonification experts or sound designers, reflects the target of the canvas: communication designers that want to include a new sensory modality in their data representation toolbox.

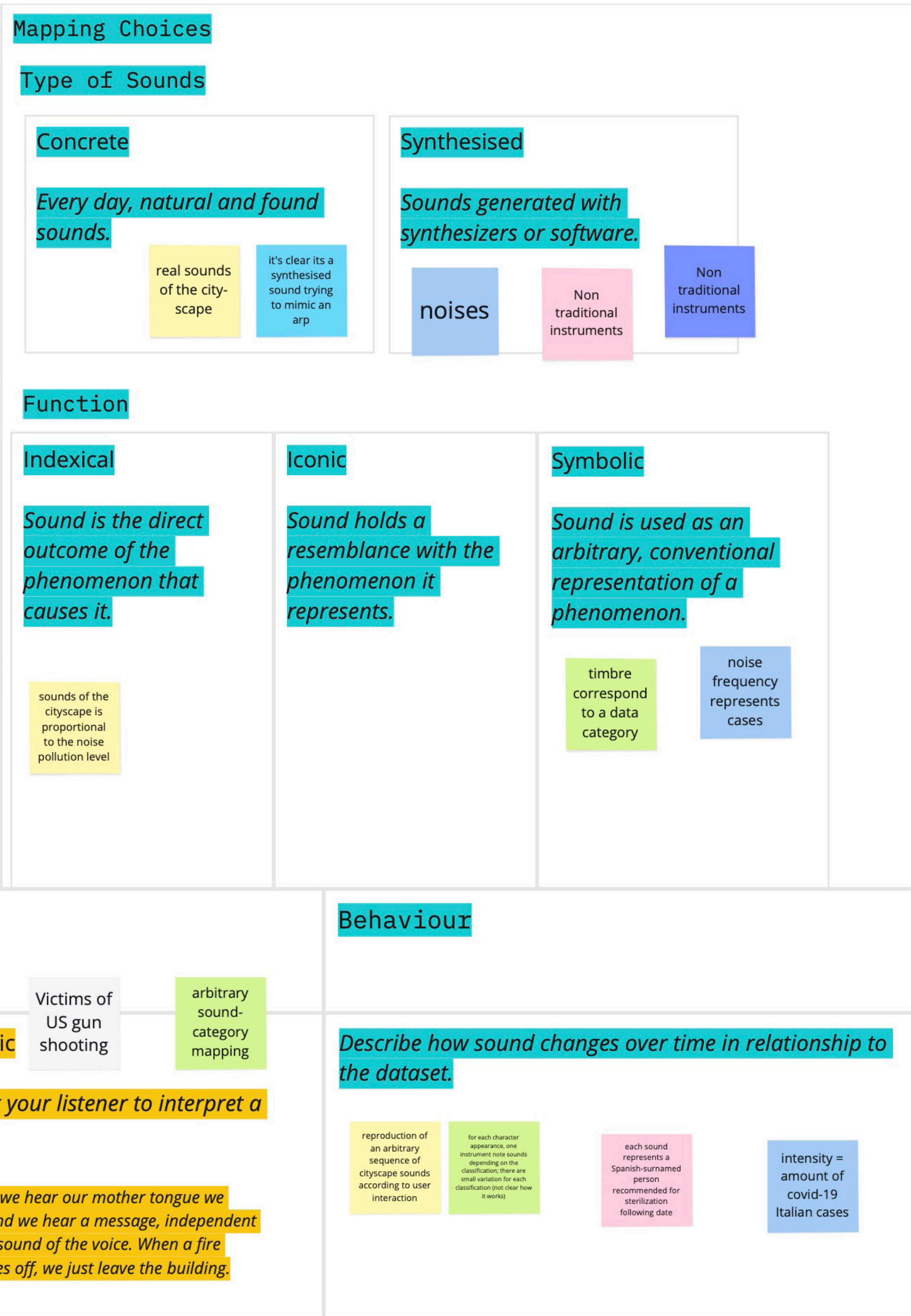
The fifteen cases were selected among those collected in the Data Sonification Archive[21] (Lenzi et al., cit.). The selection of cases was done by taking into account the availability and quality of online documentation on the design process and the sonification outcome and the variety of their characteristics in terms of use case, mapping strategy, focus and listening experience. There were no right or wrong answers I expected from the exercise. In order to have a baseline to compare the participants' choices, I had previously analyzed and categorized all the cases myself. I believed this would help me better asses the clarity and usefulness of the canvas' categories. Additionally, a group discussion was planned after the analysis in order to collect more feedback from participants. The workshop was held remotely through a combination of Zoom and Miro. After a brief introduction to the sonification canvas and to the case histories' database, the five participants were split in to two groups. Each group was accompanied by an observer who followed the analysis process and intervened if required. The two observers were Prof. Paolo Ciuccarelli and myself. A complete list of the fifteen cases, together with the information provided for each case, is available in the Appendix. As a first round of analysis, each group focused on the same cases to allow me to compare the strategy of each group. After a short break, we decided that each group would split the remaining cases so that all fifteen cases could be completed. A color-code was associated to the cases, in order to fill the canvas with colored post-its that could easily be matched to the cases' database and used to visually keep track of progress. Figure 22 shows the canvas completed by group B. Results for group A are available in the Appendix.

21. http://www.sonification.design

As shown in Fig.22, several comments were made during the group exercise and directly noted in the canvas. Other comments, together with doubts and requests for clarification, were made directly to me during the final group discussion. In general, feedback from the two groups was consistent. The participants found the canvas useful for gathering a deeper understanding of the mechanisms behind the use of sound to represent data. The most common doubts referred to the examples I chose to illustrate the blocks 'types of sounds' and 'listening experience'. Mainly, there was no shared interpretation of the examples by the participants. In some cases, the examples were heavily misunderstood, leading to a high degree of confusion about the meaning of the blocks and consequently, about the categorization of the case. Another issue was raised about the usage of what participants interpreted as specialized sound design jargon. That was the case, for example, with the words 'concrete' in the block 'type of sounds' and 'reduced', 'causal' and 'semantic' in 'listening experience'. Some of the participants appreciated the potentiality of the canvas for educating designers in the field of sound studies but highlighted that a more detailed explanation of the terms had to be provided. It was suggested that in a potential real usage of the canvas, an instruction manual could be provided along with the sonification canvas. The theme of a mono-modal or multi-modal approach to the representation emerged.

When developing the canvas, I assumed that decisions around the use of sound as the main representational modality or as a complement to visualization would emerge in the 'use case' block as part of the context. For the participants, however, this seemed to be an overarching decision that heavily influenced the design decisions related to the mapping strategy. The block 'focus', which I considered problematic during the canvas' design phase, was either left blank (group A) or marked as ambiguous (group B). During the final discussion, participants highlighted that even if the goal of the block was intuitively clear (i.e., stimulating the designer to reflect on the overall approach to the sonification) it was not evident how each case could be placed within the two categories. It was also suggested that the two poles would work as the two extremes of a continuum between an analytical and a narrative approach, between – as a participant to the workshop put it - "sonification as a tool for analysis or as a means of telling stories."

Fig. 22: group B sonification canvas for the first workshop.


Mapping Choices
Type of Sounds
Concrete
Every day, natural and found sounds.
real sounds of the city-scape
it's clear its a synthesised sound trying to mimic an arp
Synthesised
Sounds generated with synthesizers or software.
noises
Non traditional instruments
Non traditional instruments
Function
Indexical
Sound is the direct outcome of the phenomenon that causes it.
sounds of the cityscape is proportional to the noise pollution level
Iconic
Sound holds a resemblance with the phenomenon it represents.
Symbolic
Sound is used as an arbitrary, conventional representation of a phenomenon.
timbre correspond to a data category
noise frequency represents cases
Victims of US gun shooting
arbitrary sound-category mapping
your listener to interpret a
we hear our mother tongue we
d we hear a message, independent
sound of the voice. When a fire
s off, we just leave the building.
Behaviour
Describe how sound changes over time in relationship to the dataset.
reproduction of an arbitrary sequence of cityscape sounds according to user interaction
for each character appearance, one instrument note sounds depending on the classification; there are small variation for each classification (not clear how it works)
each sound represents a Spanish-surnamed person recommended for sterilization following date
intensity = amount of covid-19 Italian cases

Based on the feedback gathered during the workshop, I updated the sonification canvas as shown in Figure 23. In general, I changed the approach to the phrasing of each block, trying to make each sentence more explicit and in the form of a question. This way, I hoped to trigger a self-awareness process in which the designer, by questioning her own assumptions and work process, makes explicit the implicit decisions taken during the development of the project.

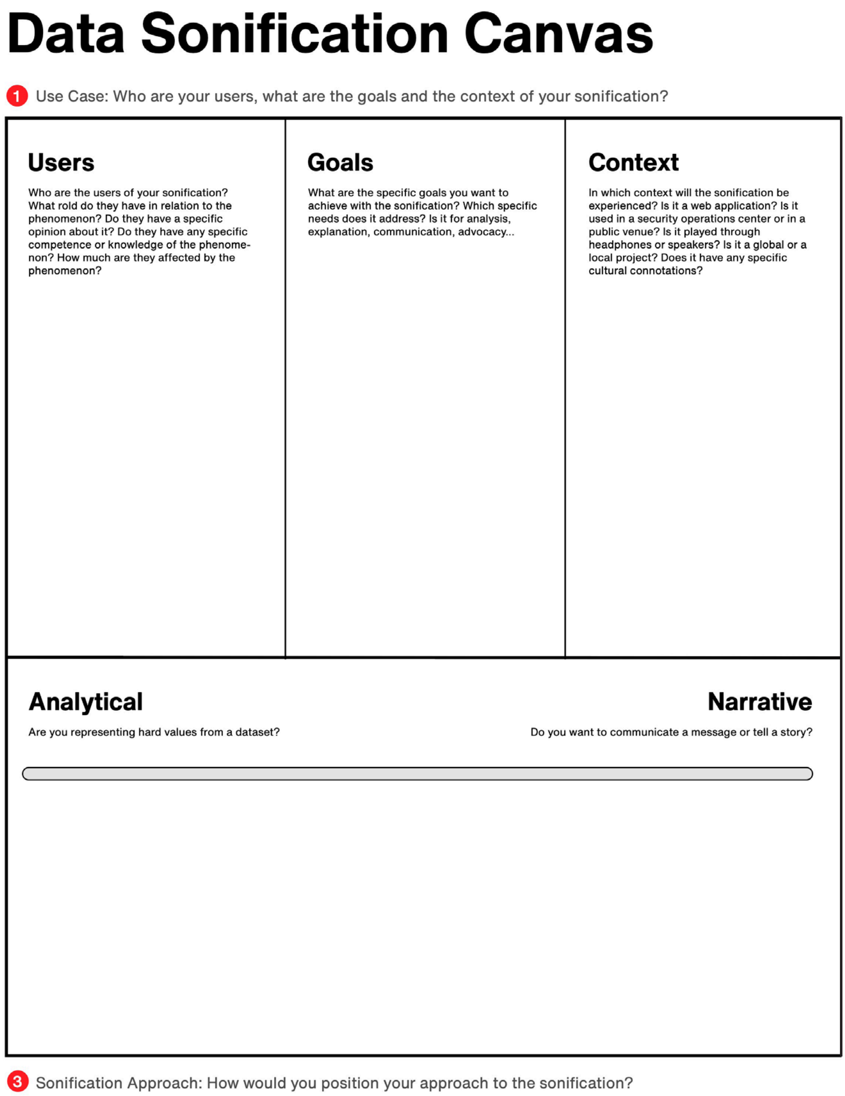


Fig. 23: data sonification canvas, v.0.2.

The overall goal of this maieutic process is to stimulate reflection and free exploration that might lead to the development of unexpected paths and an overall improvement in the design process. The canvas is meant to be used iteratively i.e., at the start of the project and then on a continuous basis during the design phase until final completion. Ideally, the canvas should keep track of the changes in the process from the conceptualization of the sonification to its delivery: it would constitute a sort of work journal as well as a documentation of the process itself.

2 Mapping Choices: How do you map data parameters to sound parameters?

Type of Sounds

Synthesised: are the sounds generated with a synthesizer? Is it intended to mimic an existing sound?
Concrete: is the sound from nature or from human activities? Is it sourced from analogue musical instruments?

Behaviour

What are the rules that link changes in the dataset to changes in the sounds?

Functions

Indexical: is sound directly produced by the phenomenon you want to represent?
Iconic: is sound similar to the phenomenon you want to represent?
Symbolic: is sound arbitrarily related to the phenomenon you want to represent?

e.g.
Indexical: the intensity of rain is detected by listening to the sound it emits.
Iconic: the intensity of rain is mimicked by the sound of rice grains falling on a surface.
Symbolic: the intensity of rain is represented by the sound of different musical instruments.

Multi-modality

Are you using only sound or is sound coulpled with other sensory modalities?

Causal

Will they gather information on the phenomenon that produced the sounds?

e.g. when you tap a container and the sound it makes give you information on how full it is.

Semantic

Will they apply a code to interpret the sounds?

e.g. when you need to apply Morse code to decipher the message contained in a sound

Reduced

Will they focus on the sound itself and its inner characteristics?

e.g. when we distinguish the interval between two notes or the pitch of a bridsong.

4 Listening Experience: How do you imagine your users will listen to the sonification?

Figure 23 illustrates other minor changes following the first workshop:

- In this second version, all the requirements were phrased as questions. Subtitles in the form of questions, were added to three macro-blocks (use case', 'mapping strategy', 'listening experience').
- In the block 'types of sound' I updated the wording and included a question to account for the usage of synthetized sounds which would mimic real-world sounds (as in the case of synthetized musical instruments, an example brought by one of the workshop participants).
- A block requiring the designer to make explicit the choices around the use of sound alone or coupled with other sensory modalities ('multi-modality') was added.
- The examples in the 'function' block were updated. To increase the consistency between the three options (indexical, iconic, symbolic), I used the same example (representing data on rain precipitation).
- In the fourth block ('focus') the title – which had led to a certain confusion - was deleted in favor of an explanatory question. In the same block, following comments during the workshop, I added a non-graded scale which is meant to represent the continuum between an analytical and a narrative approach to the representation of a phenomenon. The designer is encouraged to position her approach along the continuum and if possible elaborate on the decision.
- Finally, in the 'listening experience' block, examples were updated. Given the difficulties in finding appropriate examples that would be understood by an audience not familiar with the electroacoustic music and sound design jargon, I finally decided to adopt the examples used by Chion himself (1994 cit.) when he describes the three modes of listening.

With the goal of validating the new canvas' prototype in a real-world scenario, participants in the first workshop were asked to take part in a second workshop in which they would use the canvas as a tool for the early conceptualization of a real sonification project.

Four designers took part in the second workshop, three of whom were experts in data visualization who had participated in the first workshop, while the fourth was a sound designer with no previous knowledge of the canvas and who was not familiar with data sonification. The workshop lasted about two hours, with another thirty minutes dedicated to sharing comments, suggestions and feedback. Participants were asked to simulate the use of the sonification canvas in a real work situation. After a short introduction to the new version of the canvas, four design briefs (written by me) were shared with the participants, each one outlining a realistic project in which a client asked the designer to use data sonification. We framed the goal of the workshop as follows: the validation of the sonification canvas as a real-world design tool to support communication and information designers (used to work with data visualization) in using data sonification as an alternative or complementary modality for the representation of data. Specifically, we imagined the canvas would be used at the beginning of the work, during the phase in which designers, after receiving the client's brief, start conceptualizing the project and evaluating what elements they will have to take into account, what decisions will have to be taken and what instruments and tools (or materials and processes) they will need in order to carry out the project. Any creative activity is, to some extent, an exercise of self-reflection on and self-awareness of the decisions one is continuously asked to take.

Participants were split in two groups and each group chose two of the creative briefs. Due to time constraints, only one case was completed by each group. The cases are described as follows and links to the relevant database were provided during the workshop:

Case 1 (Group A)

Theme: introduction to fine art for visually impaired visitors (in particular, people with low vision). Using sonification to represent the catalogue of the artworks owned by the Provincia Autonoma di Trento (a region in the North of Italy) and on display at the local contemporary art museum 'MART'. The sonified catalogue will be a tool that visually impaired visitors explore prior to a visit to the physical collection.

Client: MART (Contemporary Art Museum of Trentino).

<u>Database</u>: catalogue of the artworks owned by the Provincia di Trento, available through the website of the open data project 'Trentino Open Data'.

<u>Context of usage</u>: physical space and website. The sonification could be a sound installation or a kiosk at the entrance to the museum. At a later stage, the experience could be transferred to a website.

Figure 24 shows the canvas of the first case completed by group A. Numbers (in green post-its) show the order in which the blocks were completed. The canvas also includes several comments made by this group that were later discussed during the feedback session.

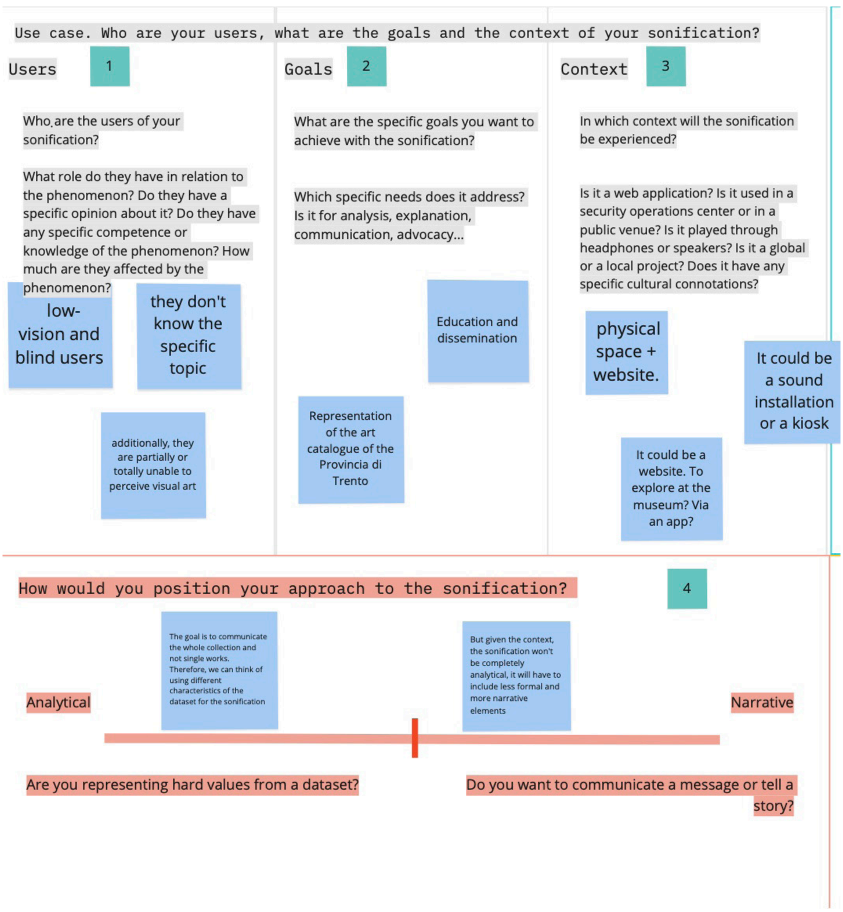
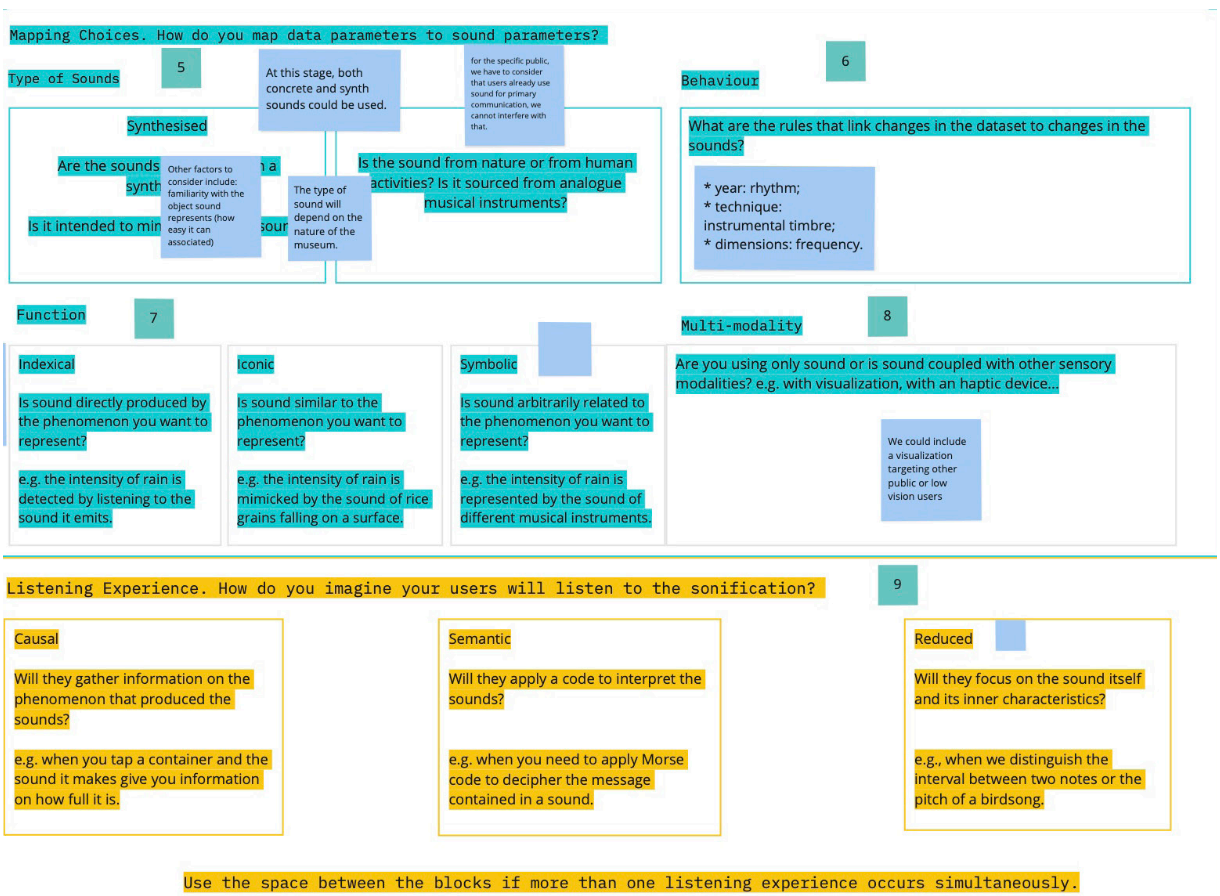


Fig. 24: sonification canvas completed by group A.


Mapping Choices. How do you map data parameters to sound parameters?
Type of Sounds
5
At this stage, both concrete and synth sounds could be used.
for the specific public, we have to consider that users already use sound for primary communication, we cannot interfere with that.
Synthesised
Are the sounds
a
synth
Is it intended to min
sour
Other factors to consider include: familiarity with the object sound represents (how easy it can associated)
The type of sound will depend on the nature of the museum.
Is the sound from nature or from human activities? Is it sourced from analogue musical instruments?
Behaviour
6
What are the rules that link changes in the dataset to changes in the sounds?
* year: rhythm;
* technique: instrumental timbre;
* dimensions: frequency.
Function
7
Indexical
Is sound directly produced by the phenomenon you want to represent?
e.g. the intensity of rain is detected by listening to the sound it emits.
Iconic
Is sound similar to the phenomenon you want to represent?
e.g. the intensity of rain is mimicked by the sound of rice grains falling on a surface.
Symbolic
Is sound arbitrarily related to the phenomenon you want to represent?
e.g. the intensity of rain is represented by the sound of different musical instruments.
Multi-modality
8
Are you using only sound or is sound coupled with other sensory modalities? e.g. with visualization, with an haptic device...
We could include a visualization targeting other public or low vision users
Listening Experience. How do you imagine your users will listen to the sonification?
9
Causal
Will they gather information on the phenomenon that produced the sounds?
e.g. when you tap a container and the sound it makes give you information on how full it is.
Semantic
Will they apply a code to interpret the sounds?
e.g. when you need to apply Morse code to decipher the message contained in a sound.
Reduced
Will they focus on the sound itself and its inner characteristics?
e.g., when we distinguish the interval between two notes or the pitch of a birdsong.
Use the space between the blocks if more than one listening experience occurs simultaneously.

Case 2 (Group B)

Theme: Public project to raise awareness on the importance of trees and the issue of global deforestation. The main goal is the sonification of the Wood Wide Web project (Mcfarlan 2016).

Client: Global Forest Initiative.

Database: provided by the Wood Wide Web project and based on the Global Forest Initiative.

Context of usage: web experience.

Figure 25 shows the canvas for this case completed by Group B. The designers used pink post-its to note down relevant information obtained from the brief, as well as comments, decisions. Blue post-its with numbers show the order in which the blocks were completed.

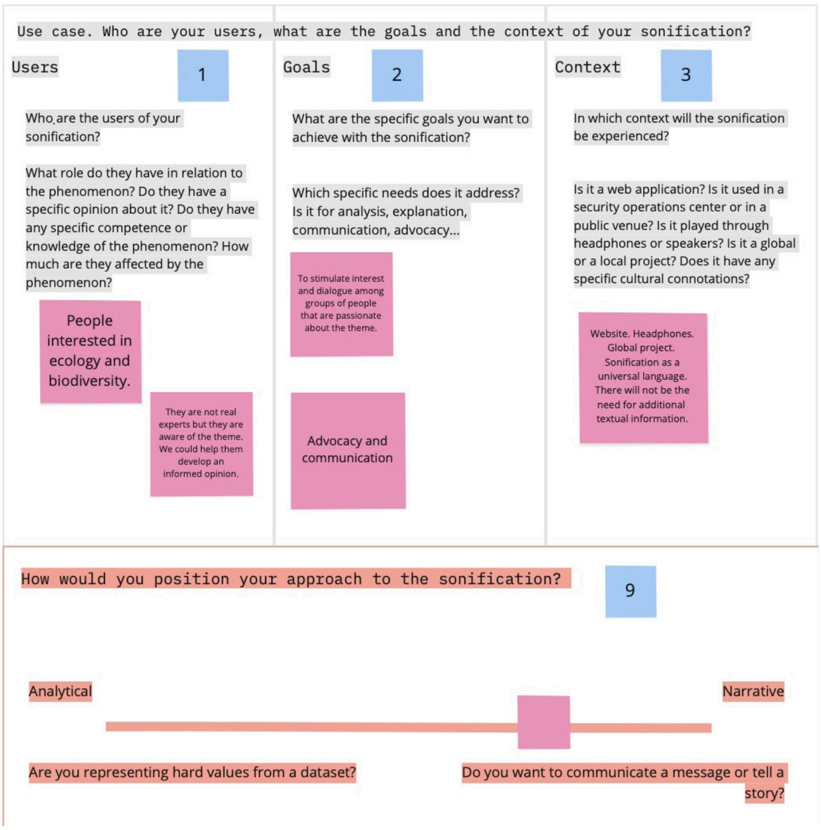
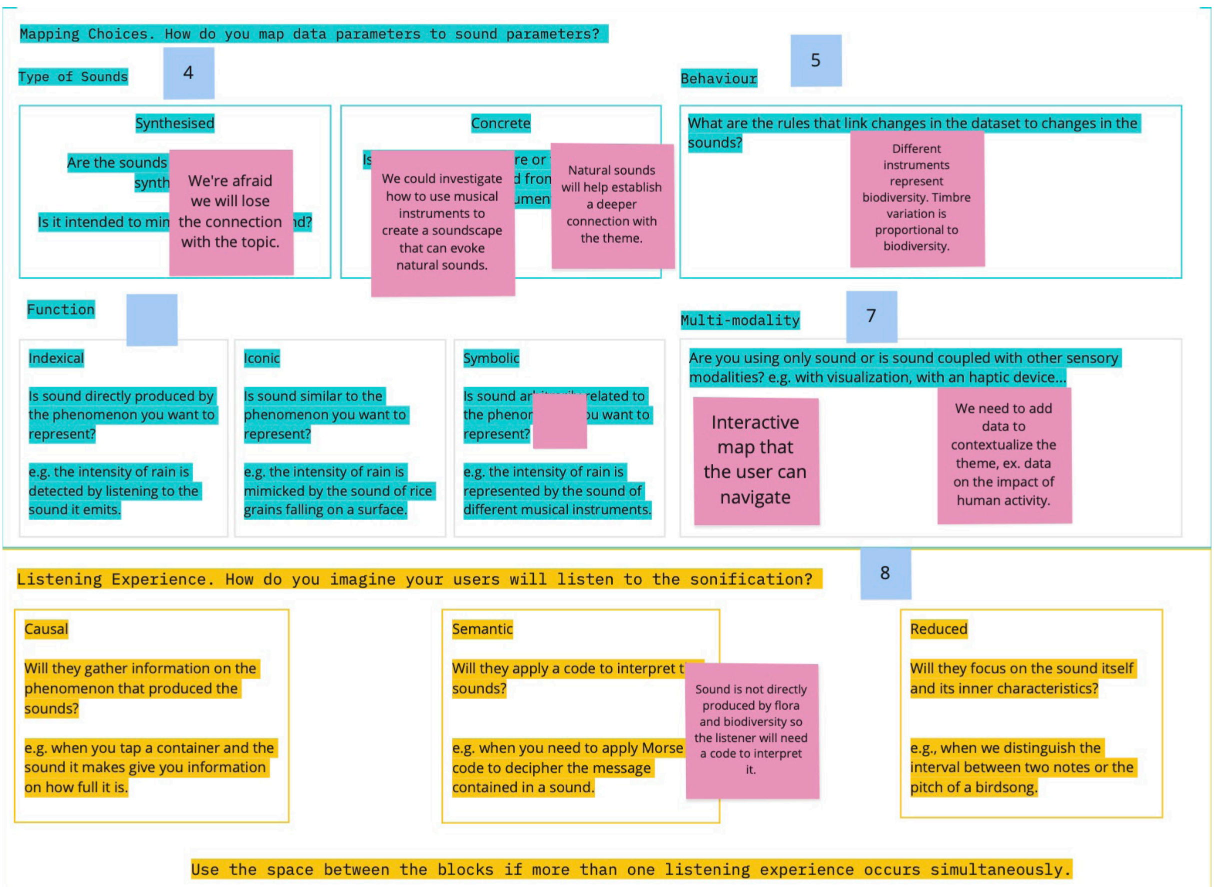


Fig. 25: sonification canvas completed by group B.



Mapping Choices. How do you map data parameters to sound parameters?
Type of Sounds
4
Synthesised
Are the sounds
synth
Is it intended to mim
nd?
We're afraid we will lose the connection with the topic.
Concrete
re or
d fron
umen
We could investigate how to use musical instruments to create a soundscape that can evoke natural sounds.
Natural sounds will help establish a deeper connection with the theme.
Behaviour
5
What are the rules that link changes in the dataset to changes in the sounds?
Different instruments represent biodiversity. Timbre variation is proportional to biodiversity.
Function
Indexical
Is sound directly produced by the phenomenon you want to represent?
e.g. the intensity of rain is detected by listening to the sound it emits.
Iconic
Is sound similar to the phenomenon you want to represent?
e.g. the intensity of rain is mimicked by the sound of rice grains falling on a surface.
Symbolic
Is sound ar
related to
the phenor
u want to
represent?
e.g. the intensity of rain is represented by the sound of different musical instruments.
Multi-modality
7
Are you using only sound or is sound coupled with other sensory modalities? e.g. with visualization, with an haptic device...
Interactive map that the user can navigate
We need to add data to contextualize the theme, ex. data on the impact of human activity.
Listening Experience. How do you imagine your users will listen to the sonification?
8
Causal
Will they gather information on the phenomenon that produced the sounds?
e.g. when you tap a container and the sound it makes give you information on how full it is.
Semantic
Will they apply a code to interpret t
sounds?
e.g. when you need to apply Morse code to decipher the message contained in a sound.
Sound is not directly produced by flora and biodiversity so the listener will need a code to interpret it.
Reduced
Will they focus on the sound itself and its inner characteristics?
e.g., when we distinguish the interval between two notes or the pitch of a birdsong.
Use the space between the blocks if more than one listening experience occurs simultaneously.

Feedback from participants can be grouped into two macro-areas. The first area refers to comments on the new version of the canvas, to which the three designers who participated to the first workshop contributed. In particular:

- The use of questions, both in the subtitles of each macro-block and in each sub-block, was judged extremely positively. One of the participants stated that questions help make things clearer while another highlighted that they stimulate self-reflection on the constraints and opportunities of the project, “instead of copying and pasting from the client’s brief”.
- In general, the new block ‘multi-modality’ was welcomed as helpful.
- The reframing of the former ‘focus’ block as a continuum between an analytical and a narrative approach was deemed very useful. Comments highlighted that it helps raise awareness on what they described as a macro decision from which the mapping choices seem to depend. In the process, some of the participants used it as a ‘checkbox’ to go back to from time to time to evaluate whether the position along the continuum had changed, in an ongoing process of self-reflection.

The second macro-area highlighted challenging elements of the canvas. In particular:

- The block ‘function’ relates to semiotic concepts that can be difficult to tackle, if previously unknown. Nonetheless, participants agreed that the effort in making explicit the design decisions in terms of the function of the chosen sounds helps you think twice about aspects that influence the final outcome.
- Likewise, the box ‘listening experience’ might require some time to be interpreted if the designer is not familiar with the definitions. Again, in order to improve planning and design, it was described as a useful exercise to take the viewpoint of the final user.

In the last part of the workshop, a more open conversation followed, where participants either shared comments they had previously reported on the canvas or engaged in a conversation around critical topics (mostly, functions and listening modes). Among the most interesting questions that emerged, the following are worth noting for future consideration:

- Can we say that when we represent a phenomenon that naturally emits sound (when it occurs in the real-world), then the designer can choose any of the three semiotic functions of sound whereas if the phenomenon has no sound attached then the sonification is necessarily symbolic?
- Is there a relationship between listening modes and the semiotic functions of sound? In particular, do all symbolic sounds require semantic listening and are all indexical sounds listened to causally? What about phenomena and objects whose sound – despite being conventional – has penetrated our habits so extensively that we almost consider them indexical (e.g., sounds emitted by electronic devices)?

One of the participants suggested, based on these considerations, re-arranging the canvas' blocks in order to suggest a correspondence between the function and listening experience, while other participants disagreed. We concluded that the repeated usage of the canvas in real-world projects could form a corpus of knowledge that would help unravel such questions and highlight potential correspondences that could lead to a review of the process (and consequently of the canvas).

As shown in Figure 24 and Figure 25, the two groups adopted different strategies when it came to the order in which the blocks were filled. Perhaps predictably, both groups started with the definition of the use case. As a second step, group A proceeded to locate their approach on the analytical/narrative continuum whereas group B started meddling with the practical decisions which the 'mapping choices' block involves. Based on their own judgement, once specific and practical decisions have been taken, choices on both listening experience and the analytical/narrative focus are inevitably constrained. On an opposite note, group A highlighted that addressing the overarching approach to the project right at the start helped them engage in an extremely relevant conversation about the explicit and implicit goals of the project. Additionally, it helped focus on nuances that they might have overlooked if they first determined the specific technical choices (in the mapping block). Consequently, some of the participants suggested the canvas came with a pre-determined order for the completion of the blocks with the goal of fostering a free exploration of the possibilities offered by the sonification. Nonetheless, in conclusion, the group seemed to agree that, especially in an early stage of the design process, many of the decisions that the canvas tries to make explicit seem to happen in the mind of the designer 'all at once'. Each of the blocks of the canvas is indeed connected to the other as they all co-exist in the designer's mind during these very first creative moments. The order in which the designer unpackages them into individual blocks could have consequences on the final artifact, but it also depends on personal preferences, work habits, artistic decisions, client requests and many other factors that would be difficult to foresee. We concluded that constraining this order would be difficult, or even counterproductive, and that every user of the canvas should be free to decide how to use this tool.

With the presentation of the current version of the sonification canvas I conclude the first part of this work. This part constitutes the theoretical investigation into the potentialities and possibilities of a design-driven approach to sonification. Results obtained from the literature review, interviews with experts, analysis of case studies and participatory activities such as workshops, are condensed in the sonification canvas as a tool that aims to support communication designers in the use of sound for the representation and communication of data. In the first chapter, I expanded on the research questions in search of a definition of sound design as an independent discipline that crosses several fields of application. The exploration brought me to also investigate work processes and experimental protocols for the design and validation of sound artifacts. In the second chapter, I turned to design as a discipline able to provide the context and the tools to structure sonification as a medium of communication in its own right. First, I turned to the criterion of 'intentionality' as the process through which the designer takes deliberate decisions to address specific needs, for a specific user, in a specific context. How much intentionality can be found in sonification was evaluated through five recent data sonification projects. In a second phase, I turned the attention to the daily practice of sonification designers. Through a series of interviews, I mapped a data sonification space in which several decisional areas emerged. These 'decisions' were used to build a sonification canvas as a design tool to guide authors in the process of using sonification to represent data. The canvas was validated through two workshop with communication designers (who were not familiar with sonification). The results of the workshops led to the definition of an updated sonification canvas which I propose as a tool for the integration of sound in data visualization.

Part II.
Sviluppo.
Data sonification for anomaly detection.

Part II. DATA SONIFICATION FOR ANOMALY DETECTION

The second part of this work illustrates two practical cases of data sonification applied to the detection of anomalies in digital-physical networks (a water distribution network) and fully digital networks (an Internet network). The two cases, which I call 'Design Actions', were developed during the course of my doctoral investigation together with international partner institutions.

The first case consisted in the design, development and testing of a series of sonification prototypes for the monitoring of cyber-attacks on the water distribution network of a model medium-sized city. It was developed in collaboration with Singapore University of Technology and Design (SUTD) and Ginevra Terenghi, a Communication Design Master's student at the Politecnico di Milano whose final dissertation was built on this case. Over a nine-month period, Terenghi and I worked on the design of two rounds of data sonification prototypes. The prototypes were used to sonify data collected from an anomaly detection algorithm which was developed by SUTD with the objective of identifying cyber-threats to water infrastructure, specifically to water distribution. Four different prototypes based on different sonification strategies, as well as different types of sounds, were tested with six experts in water networks and cyber-security. The design, development and testing of the prototypes are described in detail in Chapter 4.

Chapter 5 describes the second Design Action. Developed during a 12-month internship as member of the AI research team of the Spanish company Ibermática, it applies sonification to the real-time monitoring of anomalies in a medium-sized Internet network. Taking into account the results of the quantitative and qualitative research described in Chapter 4, a sonification prototype was designed which uses soundscape of a rainforest to alert operators on both anomalous behavior and specific cyber-attacks to the network. A fully functioning version of the prototype was developed using Max/MSP and Python, in collaboration with Ibermática's team and Prof. Damiano Meacci (Conservatorio di Firenze). The prototype was applied during a testbed organized in partnership with the Spanish research company Tecnalia which hosts a department that is fully dedicated to security in Internet networks. Despite the difficulties caused by the COVID-19 pandemic, the prototype was tested in a real-environment by experts in cyber-security and anomaly detection.

In both chapters, particular attention is paid to the description of each step of the design process and to the definition of the experimental protocol. These two aspects were of the greatest relevance in informing the research questions “Can a designerly approach help design better sonifications?” and “How do we validate a sonification project?”.

Additionally, the two design actions are at the core of the third research question which specifically addresses whether data sonification can account for the complexity of digital and digital-physical systems in order to help expert users detect and prevent anomalous behaviour in their daily work activity.

CHAPTER 4 —
Secondo tema, prima variazione.
Detecting cyber-attacks on water distribution networks.

Soundtrack: *Kraftwerk,Radioactivity.*

There seems to be a growing need for tools that can facilitate the communication between anomaly detection algorithms and human operators. Artificial Intelligence is being heavily introduced to monitor, in real-time, the behavior of digital and digital/physical systems such as electricity grids, water plants, Internet networks, industrial production and so on, with the goal of detecting anomalous behavior (mainly, but not only, due to cyber-attacks) at an early stage and making predictions on future anomalies. In the recent past, the use of sound as a substitute or as an addition to visual dashboards has attracted the interest of the research community. Examples of data sonification projects designed to support experts in monitoring tasks have covered various fields, including the continuous monitoring of medical applications in healthcare (Ballora et al. 2004); data monitoring in finance (Nesbitt and Barrass 2004); cyber-security of Internet networks (Axon et al. 2020) and IoT networks (Roddy 2018) and process monitoring in industrial production (Hermann, Hildebrandt, Langeslag, Rinderle-ma 2015). All these cases focus on a similar use case: the usage of data sonification as a complement to data visualization in order to address growing concerns on the impact of information overload (Roetlzer 2019) for human operators, based on the assumption that sound can relieve the visual channel by providing a peripheral monitoring system that does not require visual focus (Bakker, Van den Hoven, Eggen 2012). It seems that changes in acoustic patterns in peripheral sonifications are easily detected by the human ear (Vickers 2011) and would therefore attract the attention of the user while avoiding to further burden the visual channel (Ballatore, Gordon and Boone cit.). Additionally, sound composition is inherently multivariate: a series of acoustic parameters are organized in a time-based sequence where every unit synchronically handles multiple characteristics, for instance, pitch, amplitude, rhythm, timbre.

The human ear can distinguish multiple single units with individual characteristics when played at the same time for example, while listening to a musical composition or to a natural soundscape (Chion 2015, cit.). Finally, there is evidence that when used for continuous monitoring sound can enhance event prediction (Hildebrandt, Hermann, Rinderle-Ma 2014) thus helping operators prevent problems instead of reacting to emergencies when they have already occurred. Literature on auditory alarms highlights that 'intelligent' alarms should both distract operators from their main tasks while providing additional information. In particular, according to Guillaume (2011), three types of information should be conveyed by sound: firstly, an indication of how serious the failure is i.e., how urgently the situation requires the operator's attention; secondly, what caused the alarm; and thirdly, the location of the fault could be an additional useful informative layer in order to reduce the retrieval of information from the existing visualization systems.

The following sections illustrate how we applied findings from the literature to the design of a sonification system for the detection of cyber-threats to a water distribution network in the frame of the BATABAL project.

BATADAL: the BATtle of the Attack Detection ALgorithms

In recent years, water supply infrastructure has experienced a transition from a fully physical to a cyber-physical system: networked devices (smart sensors, industrial computers, telemetry units, etc.) are increasingly used for monitoring and control purposes in order to increase reliability and controllability. In an average water network of a middle-sized city, sensors are employed, for instance, to monitor the level of water in the network tanks, the pressure of water moving through pumps and the opening/closing of industrial valves which oversee and guarantee the safe transit of drinking water from the reservoirs, through the system, to our homes. Additionally, IT systems are employed to automate the activity of the whole network, for instance, by scheduling the quantity of water released from the reservoirs through the network at different times of the day or automatically turn on a pump if the level in a tank is low. Figure 26 shows a typical set-up for a so-called smart water network (Taormina et al. 2017) where sensors and programmable logic controllers are used alongside the tanks, pumps and valves. The central Supervisory Control and Data Acquisition (SCADA) system controls all the devices across the network. It is this last element that human operators in the control room use to monitor the behavior of the network at all times.

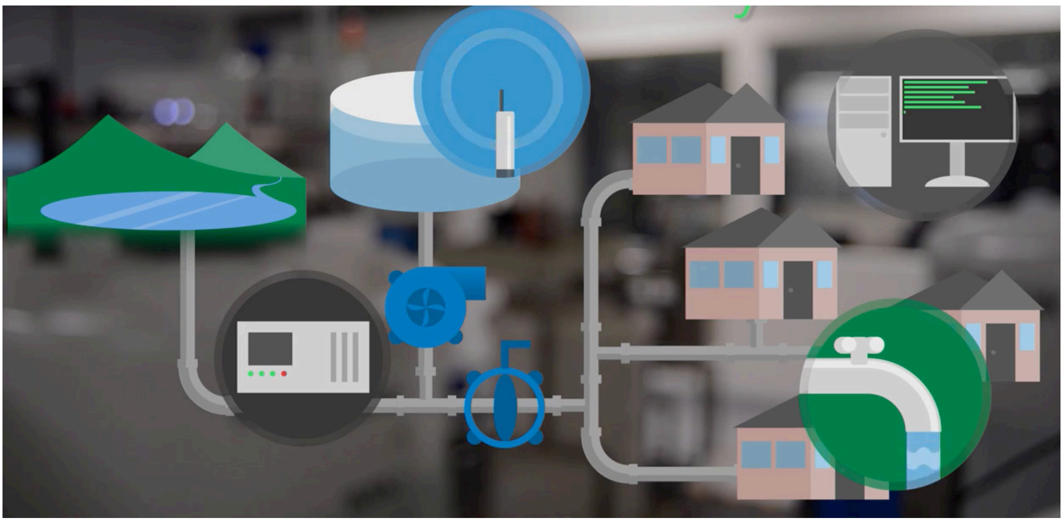


Fig. 26: Smart Water Network. Taormina et al. 2017.

Although digitalization of key infrastructures such as water plants contributes to improving efficiency and reliability and reducing water distribution costs, it simultaneously exposes cities and citizens to previously unknown forms of vulnerability such as cyber-physical attacks, "namely, the deliberate exploitation of computer systems aimed at accessing sensitive information or compromising the operations of the underlying physical system." (Taormina et al. 2018). In the past six years, the number of attacks on water infrastructure has been growing steadily: water infrastructure is the third highest targeted sector after critical manufacturing and energy (ICS-CERT 2016) Attacks can range from stealing data to damaging equipment, cutting off water supplies or even poisoning water by releasing biochemical components into the network (Taormina et al. ibid.). The modus operandi of attackers includes 'eavesdropping' attacks i.e., when the attacker manages to tap into the communication system and steal information; 'denial of service' attacks, preventing sensors from receiving data and controllers from sending instructions; 'deception' attacks, by manipulating the information sent or received by sensors; 'replay' attacks, where previously recorded information showing regular operations by the network in the SCADA system is re-played in order to hide the consequences of an attack.

Assessing the response of water distribution networks: the epanetCPA toolbox-Research centers, academics and private enterprises are multiplying their efforts to develop and deploy solutions based on Artificial Intelligence to support human operators in detecting these threats and promptly reacting, to avoid potentially critical situations that could endanger economies, local communities and even public health. In order to build reliable algorithms for detecting anomalous behavior in water networks, we need to first be able to assess the vulnerability of water infrastructures, a delicate task when we consider that real world, physical installations are located in cities and protected by the highest level of security. In the past few years, the iTrust Center for Research in Cyber Security of the Singapore University of Technology and Design (SUTD) engaged in a research project to understand how water distribution systems respond to a wide range of potential cyber-attacks. The main goal of the project is to devise solutions to prevent these attacks through a purposely designed anomaly detection algorithm. In order to do that, the team behind the project firstly explored different types of attacks through the implementation of a so-called 'attack model' i.e., a computer simulation which models different attack scenarios. Secondly, they developed a Matlab[22]- based tool called 'epanetCPA' (Figure 27) to automatically project the attack model on

22.
Matlab is a popular programming language and computing environment: https://www.mathworks.com/products/matlab.html

an ideal water distribution network created with epanet[23], the industry standard software, for modeling the hydraulic response (i.e., how water flows through distribution systems) in water distribution networks. As shown in Figure 27, the epanetCPA toolbox[24] developed by SUTD can keep track of both the physical status of the system and the digital cyber status of the system, which may have been affected by an attack.

23. https://www.epa.gov/water-research/epanet

24. The epanetCPA toolbox is available at https://github.com/rtaormina/epanetCPA.

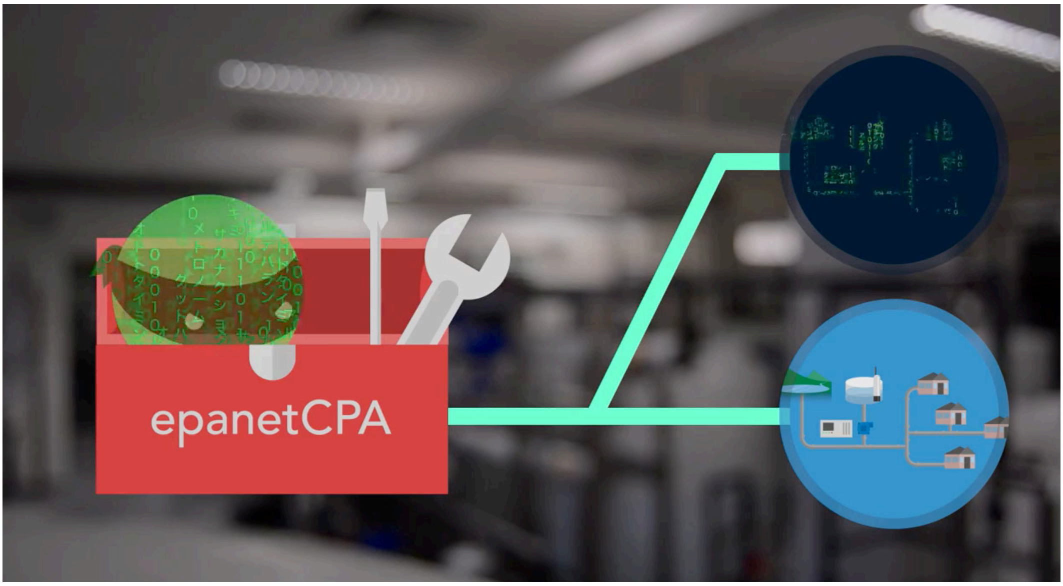


Fig. 27: diagram of the solution based on epanetCPA. Taormina et al. 2017.

Six different attack scenarios were implemented in the attack model and applied, through the epanetCPA, to the epanet-simulated water distribution system of a model town, the so-called 'C-Town' (Ostfeld et al. 2012). The C-Town water distribution system is based on a real-world medium-sized water network which includes seven tanks, eleven pumps, four valves and about 430 pipes (see Figure 28).

The six attack scenarios were applied to the C-Town water infrastructure under hundreds of different conditions which resulted in very similar outcomes (malfunctioning of water pumps, decreased level in water tanks and tank overflow). This highlighted how, despite the relative ease of spotting the consequences of an attack on the functioning of the system when a network is being monitored via the SCADA system, it might be quite difficult to identify the source of the attack i.e., the affected component, and intervene on time to preserve the integrity of the network.

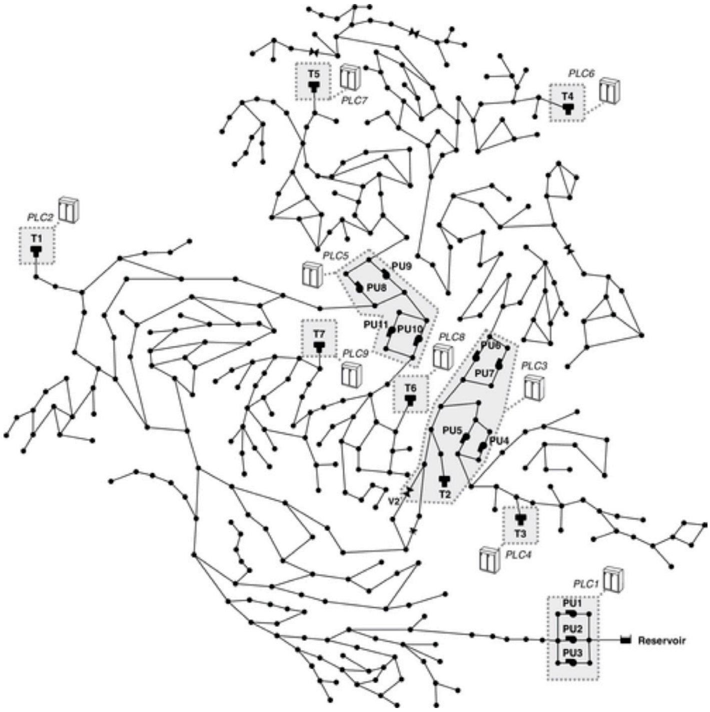


Fig. 28: C-Town.

The anomaly detection algorithm- In a second phase of the iTrust research, an international competition for the design of anomaly detection algorithms was launched in the context of the Water Distribution Systems Analysis Symposium held during the World Environmental & Water Resources Congress of 2017 (Sacramento, CA, May 21-25). The competition, known as BATADAL (the Battle of the Attack Detection Algorithms), challenged participants to develop an algorithm for the detection of cyber-physical attacks using the epanetCPA toolbox applied on the C-Town water infrastructure system. As mentioned, one of the goals of the previous phase was to assess the response of water distribution systems to cyber-threats in order to devise guidelines for the design of more flexible and resilient infrastructures. The goal of BATADAL was, on the other hand, to devise solutions for the early detection of all unforeseeable threats for which the systems cannot yet be prepared, "without issuing false alarms" (Taormina et al. 2018 cit., p.4). To recap, the main technical challenges posed to participants were: to disclose the presence of an ongoing attack in the minimum possible time to avoid issuing false alarms, and as an optional feature, identify which components of the system had been compromised.

A use case for sonification- As the BATADAL challenge requirements highlight, anomaly detection algorithms are still exposed to a certain degree of error which may result in issuing false alarms. False alarms could, in turn, (mis)lead operators into taking the wrong decision in a moment of danger. Additionally, even if the different cyber-attacks tend to provoke the same response in the network, the affected digital component might be different. Identifying the affected component (and communicating it to the operator) is therefore crucial yet "a non-negligible challenge in large water networks" (Taormina et al. 2018 cit., p.5). Thirdly, all digital components of the network can be affected by an attack, including the central SCADA system and its information could be, consequently, non-reliable. A separate monitoring system - streaming data from the anomaly detection algorithm and thus bypassing the SCADA system - should be designed in order to provide the operator with an additional source of control.

At the same time, SOCs are already heavily reliant on visual screening for the storage, management and analysis of incoming data. Adding yet another visual interface dedicated to the real-time monitoring of cyber-threats would confront analysts with a higher risk of information overload. Additionally, during their normal working day operators of a water plant are expected to carry out other (again, mainly visual) tasks such as reading and writing reports, analyzing historical data sets, answering the occasional phone call, interacting with colleagues. In early 2018, a conversation started between SUTD and myself to explore the use of data sonification as a complement to visualization in order to communicate alerts on cyber-attacks which could also help prevent false positive, bypass the SCADA system and allow the operator locate the affected component.

Data sonification design - Research

The first Design Action was developed between October 2018 and April 2019. The case was also the subject of the master's in Communication Design thesis of Ginevra Terenghi, which I co-supervised. The case was awarded the Best Use of Sound Award at the 25th International Conference on Auditory Display (ICAD 2019) at Northumbria University (23-27 June 2019). In this case, we used sound as the main sensory modality to peripherally monitor cyber-threats to the water infrastructure network of C-Town.

The case uses data simulated for the BATADAL challenge: "The goal of the battle was to compare the performance of algorithms for the detection of cyber-physical attacks, whose frequency increased in the past 10 few years along with the adoption of smart water technologies. The design challenge was set for C-Town network, a real-world, medium-sized water distribution system operated through Programmable Logic Controllers and a Supervisory Control And Data Acquisition (SCADA) system. Participants were provided with datasets containing (simulated) SCADA observations, and challenged with the design of an attack detection algorithm." (Taormina et al. cit. 2018). Seven teams, from both academia and industry, took part in the challenge of contributing in the design of a novel solution for the identification of cyber-attacks to the water infrastructure of C-Town. Each solution was evaluated based on the time needed to identify the attack, the avoidance of false alarms and the identification of the system component that was compromised.

C-Town- C-Town is a medium-size model town with a water infrastructure simulated via epanet, a Matlab-based software commonly used in engineering simulations to assess the response of water infrastructures during both normal operations and accidents. The water network of C-Town is based on a real-world infrastructure inspired by the city of Haifa in Israel, and it has first been introduced by Ostfeld et al. (2012) for the 'Battle of the Water Calibration Network'. It is characterized by a network of 429 pipes, 388 junctions, 7 storage tanks, 11 pumps (distributed across 5 pumping 102 stations), 5 valves, and a single reservoir. Additionally, 9 PLCs (Programmable Logic Controllers) were added in order to simulate the digital additions to the physical network, and located in the proximity of water pumps, tanks and valves. The cyber network is controlled by a central SCADA system. As Figure 29 shows, the whole network is divided into five districts which correspond to districts in the city of C-Town. Water consumption in the city is assumed to be fairly regular during the whole year.

To recap, during BATADAL a series of cyber-attacks to the water infrastructure, mainly based on deception attacks, were simulated using the epanetCPA. SCADA readings were also altered through replay attacks. The performance of the algorithms was evaluated based on the time taken to detect the attack, the accuracy of the detection (i.e., the avoidance of false positives) and, as an option, the capacity to locate the specific component affected within the whole network. Incidentally, BATADAL was won by the only algorithm which adopted a model-based approach i.e., simulating the performance of the physical network under regular conditions via epanet and subsequently monitoring SCADA readings in order to evaluate the potential discrepancies between the actual and the ideal predicted performance. This approach, which differs from the most common approach of data-driven detection, is fairly similar to the one adopted in the anomaly detection of fully digital networks that will form the basis of the second Design Action (Chapter 5).

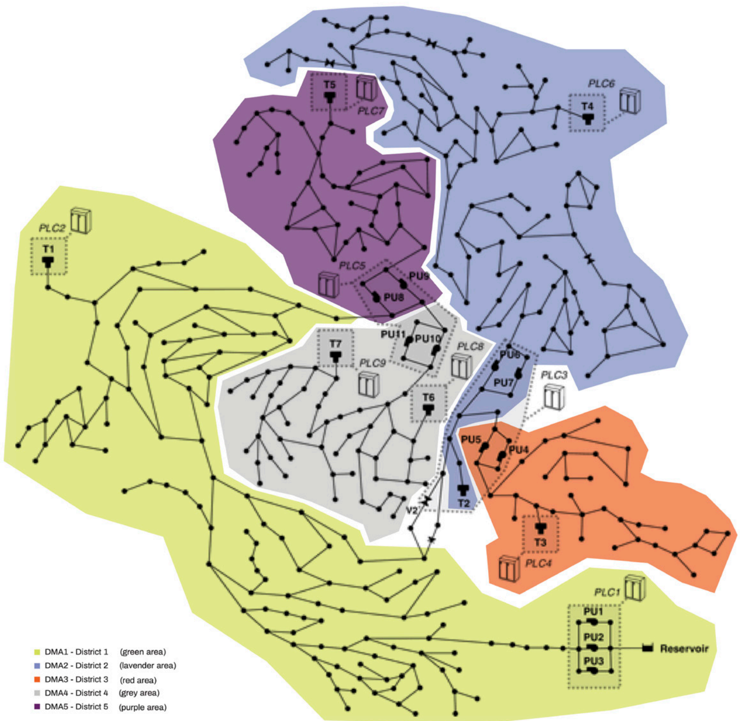


Fig. 29: map of C-Town with the five districts.

Characteristics of the data set- Participants in BATADAL were provided with three simulated SCADA datasets based on the water network of C-Town. A limitation of these types of study, recognized by the team at iTrust is, in fact, the absence of information on real-world cyber-attacks to water infrastructure. Due to national security, real datasets of compromised SCADA systems which involved cyber-attacks are not accessible to researchers hence datasets showing correlations between the attacks and the cyber-response of the system, including the SCADA database, have to be simulated.

The three BATADAL datasets included:

- Training Dataset 1: a database of 365 days of readings of the regular activity of the network, without any cyber-attacks. This dataset was used to study the behavior of the water distribution network under normal conditions.
- Training Dataset 2: a database containing seven different attacks across seven months. One attack was fully disclosed to the participants while others were only partially disclosed. The database simulates the data available to experts during a forensic investigation, where some SCADA readings would reveal details on the intrusion while others would be concealed by replay attacks.
- Test Dataset: it contained seven additional attacks over three months, none of which was revealed to the participants. It is the dataset used to test the performance of the anomaly detection algorithm.

The hydraulic time stamp (i.e., the reading of data coming from the logic controller of the physical system) of all the datasets was fixed at 15 minutes, while SCADA readings were updated every hour.

For the sonification prototype, we decided to use the Test Dataset (which is the most complete, as it was used to build the algorithms of BATADAL), cross-checked with the complete list of attacks which the iTrust team disclosed to us. Table 1 shows an excerpt of the Test Dataset, containing both regular data and attacks.

Attacks are reported in the ‘flag’ column. Data on attacks are binary i.e., the value can either be ‘zero’ (no attack) or ‘one’ (attack). The following column reports the elements of the network that were affected by the attack and the respective parameters of the element. For instance, ‘L_T3’ represents the level of tank 3 (i.e., the tank located in District 3, see Figure 29); ‘F_PU4’ indicates the flow of pump number 4 whereas ‘S_PU4’ represents the status of pump 4. Pumps are characterized by two parameters: the quantity of water in the pump (flow) and the status of the pump (on/off). As shown in the remaining columns, tanks are characterized only by the level, whereas valves, like pumps, are characterized by both flow and status.

DATETIME	flag_attack	attacked_device	L_T1	L_T3	F_PU4	S_PU4
04/01/17 00	0		0,0022094	0,002845701	0,005289089	6,80E-07
04/01/17 01	0		0,0012413	0,005563898	0,005257677	1,34E-06
04/01/17 02	0		0,0023347	0,00091524	0,004753576	5,25E-06
04/01/17 03	0		3,86E-06	0,005944347	0,00594333	9,01E-07
04/01/17 04	0		0,0020486	0,001420044	0,009540024	1,58E-06
04/01/17 05	0		0,0006754	1,43E-05	0,005672572	1,65E-06
04/01/17 06	0		0,0008609	5,05E-05	0,004276488	1,40E-06
04/01/17 07	0		0,0028127	0,000123438	0,009290928	7,12E-07
16/01/17 09	1	L_T3, F_PU4, F_PU5, S_PU4, S_F	0,0018519	0,000289975	0,00021629	1,67E-06
16/01/17 10	1	L_T3, F_PU4, F_PU5, S_PU4, S_F	0,0019533	0,00019851	0,000371092	1,59E-06
16/01/17 11	1	L_T3, F_PU4, F_PU5, S_PU4, S_F	0,0011944	0,005221615	0,011932357	8,16E-07
16/01/17 12	1	L_T3, F_PU4, F_PU5, S_PU4, S_F	7,66E-05	0,000254381	0,003937702	4,49E-07
16/01/17 13	1	L_T3, F_PU4, F_PU5, S_PU4, S_F	0,0013461	0,000901605	0,000633911	2,26E-06
16/01/17 14	1	L_T3, F_PU4, F_PU5, S_PU4, S_F	0,0009282	4,73E-05	0,000897583	6,51E-07

Table 1: test dataset.

After defining the use case for using sonification in the context of the BATADAL project, we used the test dataset to draft a first series of mapping strategies and run a first round of exploration on types of sound which will be described in detail in the following sections.

Definition of the use case- Between August 2018 and April 2019 we conceptualized, designed, prototyped and tested in a real environment a sonification prototype to represent data which simulates both the cyber-attacks and the physical response of the network water distribution network of the medium-sized model town C-Town. We imagined the sonification as a complement, and not a substitute, for the data visualization currently displayed in the central SCADA systems. Figure 30 shows a typical implementation of the visualization of a water and wastewater network in a SCADA system.

Tank, tank levels, pumps and valves with their respective flow and status can be seen on the network's map together with individual boxes indicating the exact numerical value of each parameter. The map also reports key references to locate the network component within the city (for instance, County Road, Eliza Street).

In a normal every-day situation, the personnel of the water network SOC (two operators, for a medium-sized network) monitors the SCADA system, which receives data from sensors and PLCs located near the network's components.

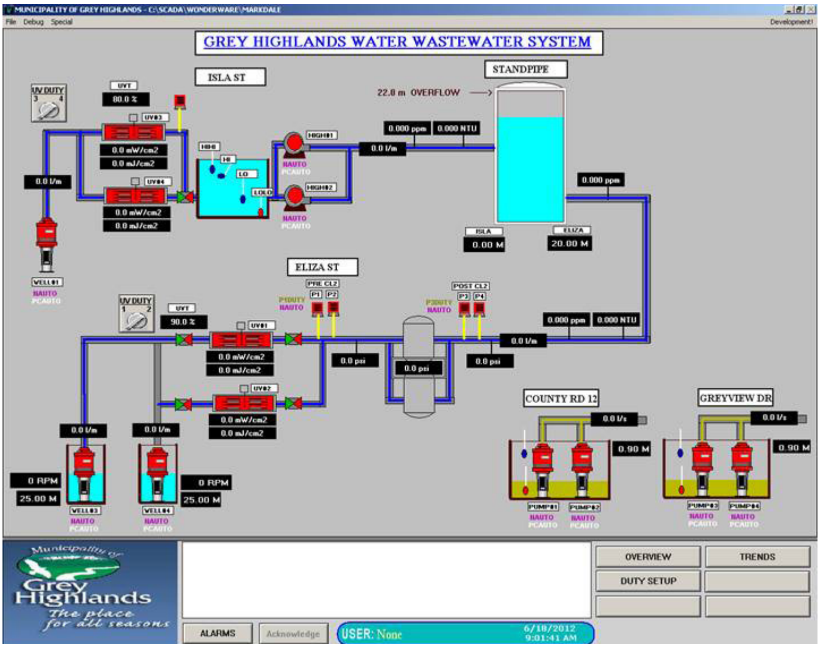


Fig. 30: SCADA system for the city of Grey Highlands by Datasoft Software Solutions.

Fig. 31: Security Operation Center setting. Courtesy of Pacific Northwest National Laboratory.

As Figure 31 shows, a SOC is typically filled with screens. In a medium-sized Center, about 10 screens are monitored in real-time by the operators: some show data from the SCADA system, some are used to assess specific trend lines or for the analysis of historical data. In our simulation, sensors and PLCs on the network are updated every fifteen minutes, whereas the SCADA system receives data every hour. In a real-world scenario, an anomaly would trigger a series of steps to alert the operator. These steps include a visual alarm (with changing colors on the SCADA visualization) as well as an acoustic alarm (mainly a traditional buzzer or siren). Once alerted, the operator would run several checks in order to locate the anomaly in the network, for example by accessing specific trend lines on screen or even requesting a manual check on the specific component on site. As an operator of city water infrastructure explained during a personal conversation with members of iTrust, currently operators tend to switch off the acoustic alarm systems. Indeed, these alarms can be particularly intrusive and since the rate of false positives in anomaly detection systems is still high the alarms often cause great disruption to the routine of the water plant staff.

We decided that the sonification would be designed with the following goals and requirements in mind:

Goal 1: The sonification should release cognitive overload from the operator's visual channel, while supporting real-time monitoring of the network.

Requirement 1: The sonification should complement, and not replace the existing visualization in order to seamlessly integrate with the well-known work environment. At the same time, it should minimize the need for yet another visualization system dedicated only to cyber-attacks. To do so, the sonification should convey just enough data to facilitate the retrieval of additional information in the existing SCADA system.

Requirement 2: The sonification should run as a continuous, peripheral monitoring system thus not intruding on the existing daily routine of the operator while being able to move to the center of the operator's attention if needed.

Goal 2: To recap, the main goals of an operator in a context of potential cyber-attacks are: to be able to identify when an attack is taking place; to identify which component is compromised; to distinguish genuine alarms from false positives.

Requirement 3: To minimize cognitive and information overload, the sonification should represent only two elements: the anomaly and the affected component.

Requirement 4: To limit the impact of false positives, the sonification should not issue a binary alarm (attack yes/attack no) but rather, it should represent anomalies in the network as a continuum: the discrepancy between the expected behavior and the real behavior of the system at a given time (what we will later call the 'reconstruction error') is represented as is, without a prior judgement on whether or not the system is deemed by the algorithm to be under attack. To interpret the gravity of the anomaly and take action, the operator would therefore rely on his knowledge of the network rather than on an Artificial Intelligence decision.

Demo. Exploring mapping and types of sounds- We ran a first round of explorations of mapping strategies and types of sounds. With the test dataset in mind, we decided each sound would represent each category of the network components (tanks, pumps and valves) while a characteristic of the sound would represent the associated variable (level, flow and status). Anomalous behavior in a specific component and variable would be represented by a progressive distortion of the sound. For the sake of the demo, we normalized the anomaly level and scaled it along a five-step scale from 1 (lowest anomaly level) to five (highest anomaly level).

Two demo sonifications were produced and shared with the iTrust team, one based on concrete sounds and the other using synthesis sounds. Sound samples were produced manually and edited using Nuendo 5 by Native Instruments. The design of the sounds was based on the following guiding principles:

- We assumed that a successful sound for peripheral monitoring would be one the operator would easily relate to real-world experiences, thus minimizing the need for training and the cognitive load while listening to the sonification. Literature on auditory alarms confirms that auditory icons - sounds with a strong indexical connection with the object they represent - seem to be promising for attracting attention as well as minimizing cognitive load and training (Guillaume cit.). These types of sound would also carry additional contextual information that do not represent data but rather contribute to the interpretation of the sounds (Walker and Kramer 2005 cit.).

We designed several sounds considering how a real tank, pump, or valve would sound.

- In the demo based on synthesis sounds we added a second layer to represent the specific variable of each component. In this case, the auditory representation of the variable would be more critical for components which can present more than one variable: for instance pumps and valves (that have flow and status.)

An excerpt of this approach can be listened to here.

- Finally, we explored the usage of metaphors to represent anomalous behavior. Specifically, we hypothesized that a distortion in the sound would alert the listener to a distortion in the data. The usage of distortion has been recently connected to the representation of uncertainty in geographical data (Ballatore, Gordon, Boone cit.), an additional experimental result that we believed supported our approach. Moreover, conveying a sense of 'wrongness' through distortion is coherent with a typical approach in sound design for film and gaming, where distortion is often used to amplify the spectator's feeling of uneasiness or an alteration in the narrative (Hilmann and Pauletto cit.).

Different digital distortion methods were explored including pitch shift, delay, filtering, addition of clicks and noise.

Lastly, we needed to locate each component in the network so that the operator could correctly identify it. We decided to play each component in a sequence representing a virtual clockwise movement through the network (see Figure 32) so that, for instance, the first tank sound heard by the operator would be the first tank at the far left of the network map.

In terms of the duration of the sound representing each component, we did not have any clues from the dataset. We therefore took a decision and tried to balance between the need to understand the information and the efficiency of the sonification. Several versions were made until we obtained a soundtrack of about 2 minutes, that played every hour to update the operator on the functioning of the whole network.

Concrete amd Abstract sounds version of the demo can be found here:
https://www.saralenzi.com/design-actions-material PSW: SonificationDesign2021

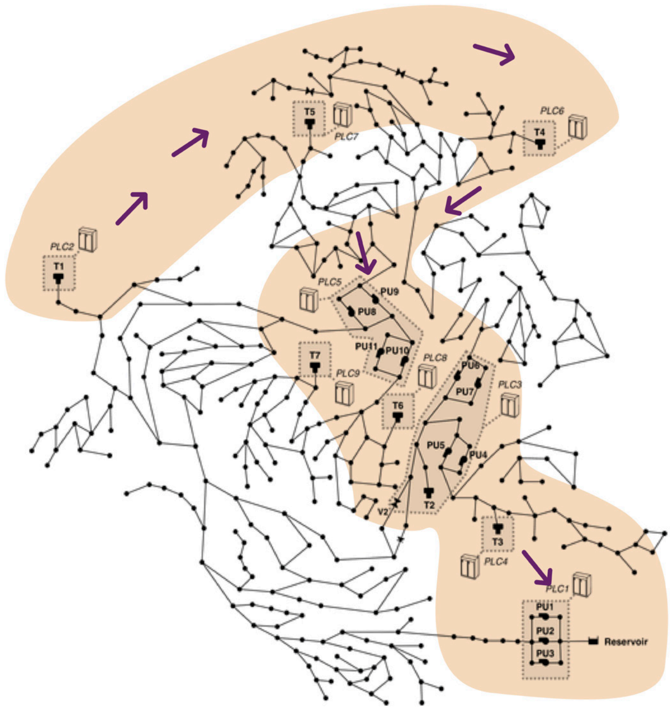


Figure 32: sequence of the sonification through the network. Terenghi 2019.

This first round of demos was shared with the iTrust team during a feedback session which highlighted several criticalities. Specifically:

- The overall duration of each instance of the sonification (about two minutes every one hour) was deemed way too long to be sustainable. In a real-world situation, the operator would have to suspend her task to carefully listen to a two-minute 'soundtrack': that was unrealistic. The operator would also have to remember the order of the components on the map while listening - a task that would be hard to fulfill even for an expert listener.
- There was an evident overload in the amount of information conveyed by the sound, which included: the type of component; the type of variable for each component; the amount of anomaly for each component and variable; the geographical location of the component. This amount of information made the prototype virtually impossible to scale up to larger networks.

Nonetheless, feedback from the session helped us pivot the strategy before the prototyping phase. In particular:

- The duration had to be radically shrunk in order to limit the impact of sonification on the routine of the control room.

- Information to be conveyed by sound had to be drastically reduced, too. In particular, while geographical information on the specific district under attack seemed very relevant, no added value seemed to come from information on components and variables as this could be easily located on the existing visualization tools if needed.
- As mentioned, an artificial scale from 1 to 5 had been introduced to represent the anomaly level of each component. The scale, which was not included in the original dataset, was pre-determined by us. As we found no clear added value in pre-determining the anomaly level, we decided to leave it to the operator to estimate the gravity of the anomaly based on her own experience.

With all these considerations in mind, we moved forward to the prototyping phase.

Data sonification design – Prototyping

We started the prototyping phase by sketching, from scratch, four new data-to-mapping options that we called 'scenarios'. The experts involved in the demo session felt that sound should convey little but clear information on the global status of the network and the gravity of the anomaly, rather than representing the status of each component. Additionally, they pointed at the possibility to locate the information on a map of the city as a relevant added value. I remind the reader that the sonification was meant to complement the SCADA visualization, which is organized geographically. If the sonification could also convey information on the geographical location of the anomaly, we hypothesized, the operator would retrieve analytical information from the SCADA more efficiently.

In a first round of prototypes, we drafted four scenarios with four different mapping strategies and four different sound types. The scenarios were shared with the project's team to gather expert feedback. Only two mapping strategies, corresponding to scenarios 2 and 3, were selected for further development. As for the types of sound, given the absence of clear guidelines from existing literature on the performance of tuned sounds versus non-tuned sounds in the context of process monitoring, we decided to develop each scenario in two versions: version A, using musical (tuned) sounds; version B, using concrete (non-tuned) sound samples. Figure 33 recaps the demo and prototypes process from the demo to the final prototypes.

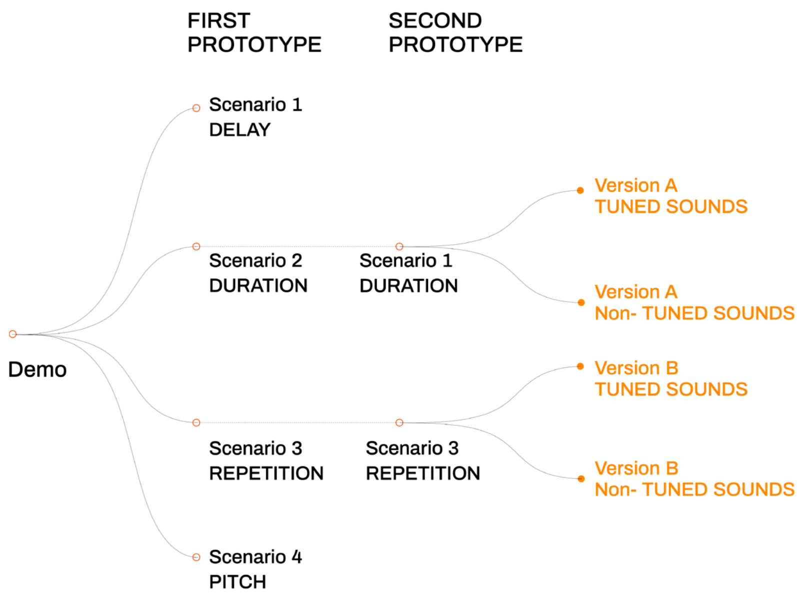


Figure 33. Data sonification prototyping phases. Adapted from Terenghi 2019.

The dataset- Data from the Test Dataset (see Table 1) were clustered by district. We then extracted the 'reconstruction error' i.e., the value of the discrepancy between the expected behavior of the network as predicted by the model and the real behavior of the network as read by the SCADA system at a given time. Table 2 shows an excerpt of the dataset. For each district we selected the maximum value of the reconstruction error independently from the component. Every hour the operator received information from the sonification on whether the network presented anomalous behavior, and, if so, in which district and what the value of the anomaly was. No information was provided on the affected component or the affected variable.

Each scenario was based on five different sounds, with each of these representing one of the districts (DMA, District Metered Area). The behavior of the five sounds over time was determined by the value of the reconstruction error, that was normalized on a scale from 1 to 10.

DATETIME	flag_attack	attacked_device	DMA-1	DMA-2	DMA-3	DMA-4	DMA-5
15/01/17 23	0		0,016398	0,0063	0,001038	0,002631	0,002787
16/01/17 00	0		0,014241	0,008878	0,001998	0,002571	0,002667
16/01/17 01	0		0,015919	0,010276	0,005453	0,001951	0,007562
16/01/17 02	0		0,017532	0,004732	0,004973	0,002587	0,006203
16/01/17 03	0		0,011046	0,00992	0,00592	0,002633	0,009551
16/01/17 04	0		0,006947	0,017489	0,00621	0,001887	0,003558
16/01/17 05	0		0,008193	0,019067	0,004349	0,002396	0,001707
16/01/17 06	0		0,006503	0,007699	0,005624	0,004218	0,002705
16/01/17 07	0		0,002881	0,002904	0,000445	0,005157	0,011893
16/01/17 08	0		0,003815	0,007031	0,001797	0,001919	0,000864
16/01/17 09	1	L_T3, F_PU4, F_PU5, S_PU4, S_F	0,002848	0,013357	0,004413	0,001835	0,000986
16/01/17 10	1	L_T3, F_PU4, F_PU5, S_PU4, S_F	0,002042	0,008338	0,005063	0,002602	0,004019
16/01/17 11	1	L_T3, F_PU4, F_PU5, S_PU4, S_F	0,011932	0,093463	0,154485	0,008798	0,002956

Table 2. Database with the reconstruction error for each district.

The four scenarios- The four scenarios followed the same general mapping concept. Every hour the sonification played all the sounds - one for each district - as a short update on the status of the network. In order not to disrupt the operator’s routine we kept the sonification as short as possible, with each individual sound lasting a maximum of one second. In all scenarios, in a regular situation, the five different sounds for the five districts play synchronically every hour for a maximum of one second. In the event of an anomaly, the behavior of each sound would be altered differently based on each scenario:

In scenario 1 - Delay, the sound of a district presenting anomalies would play later than the others, proportionally to the value of reconstruction error.

In scenario 2 - Duration, the sound of the anomalous district would last longer, proportionally to the value of reconstruction error.

In scenario 3 – Repetition, the sound of the anomalous district would be repeated for over ten seconds. The frequency of the repetition would increase proportionally to the value of the reconstruction error.

In scenario 4 - Pitch, we introduced a variant. The five sounds would play in a sequence, starting from DMA1. All the five districts’ sounds would initially have the same pitch. The pitch would increase proportionally to the value of the reconstruction error.

At the following links, the reader can listen to an example of each scenario in a regular situation, a medium-anomaly and a high-anomaly while following the visual score: https://www.saralenzi.com/design-actions-material (PSW: SonificationDesign2021).

Scenario 1 – Delay

In a regular situation, the five sounds of scenario 1 – Delay, would play in synchronous, for instance, at 10:00AM:

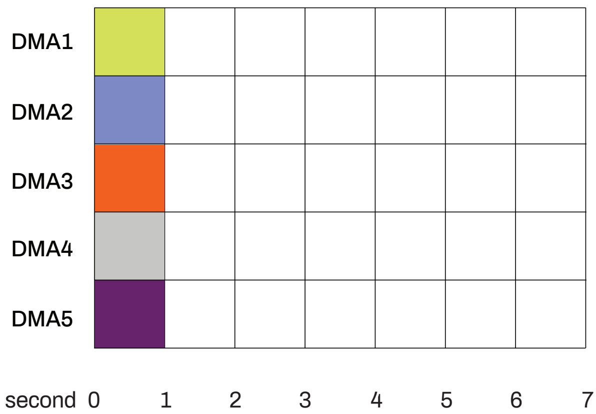


Fig. 34. Scenario 1, regular situation. Terenghi 2019.

Sound example can be found at:
https://www.saralenzi.com/design-actions-material PSW: SonificationDesign2021

At 11:00AM, at the appearance of some anomalous behavior in DMA 2 and DMA3, the corresponding sounds will play shortly after the others:

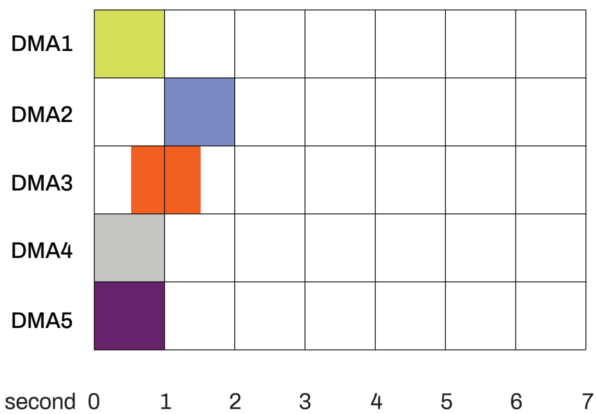


Fig. 35. Scenario 1, anomalous situation. Terenghi 2019.

Sound example can be found at:
https://www.saralenzi.com/design-actions-material PSW: SonificationDesign2021

At 12:00PM, as the attack progressively affects the network, all the districts are played with different delay times: for instance, DMA1 will be heard at 12:04.000, DMA2 at 12:01.350, DMA5 at 12:04.485.

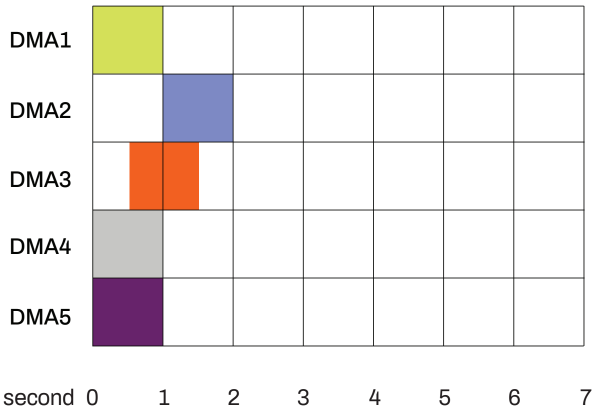


Fig. 36. Scenario 1, highly anomalous situation. Terenghi 2019.

Sound example can be found at:
https://www.saralenzi.com/design-actions-material PSW: SonificationDesign2021

Scenario 2 – Duration

In a regular situation, the five sounds of Scenario 2 – Duration, would play in synchronous at 10:00AM with the same duration of one second.

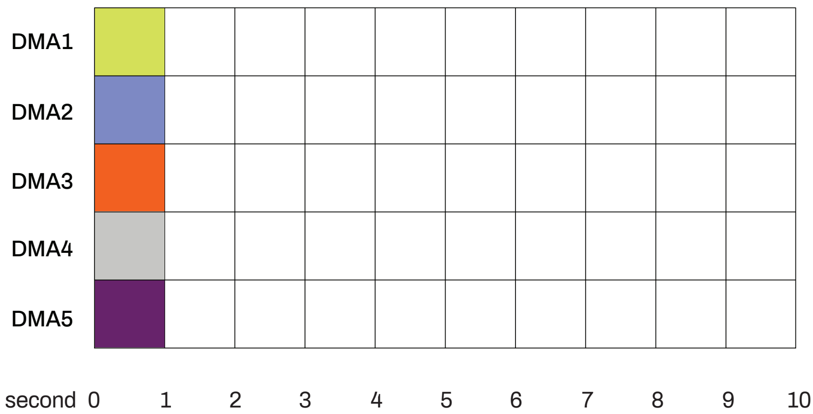


Fig. 37. Scenario 2, regular situation. Terenghi 2019.

Sound example can be found at:
https://www.saralenzi.com/design-actions-material PSW: SonificationDesign2021

At 11:00AM, at the appearance of some anomalous behavior in DMA 2 and DMA3, the corresponding sounds will play longer than the others:

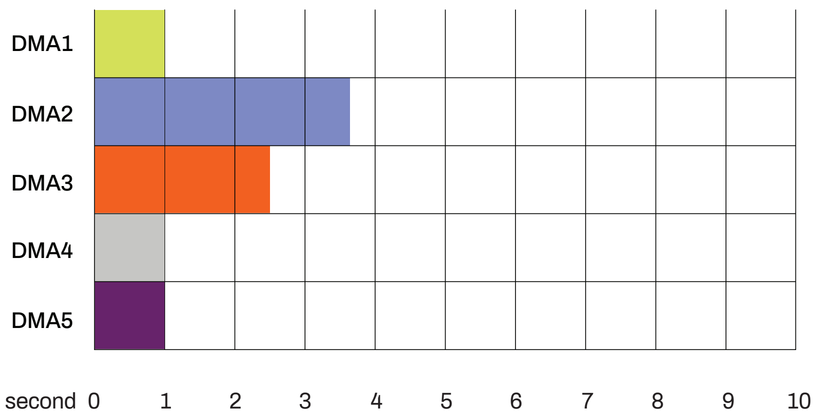


Fig. 38. Scenario 2, anomalous situation. Terenghi 2019.

Sound example can be found at:
https://www.saralenzi.com/design-actions-material PSW: SonificationDesign2021

At 12:00PM, as the attack progressively affects the network, all the districts are played for longer than at 10AM: for instance, DMA1 will be heard until 12:05.000, DMA2 until 12:02.250, DMA5 at 12:08.275.

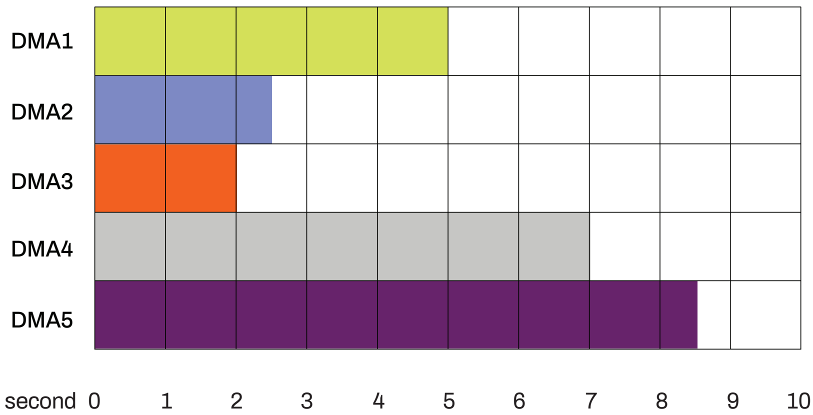


Fig. 39. Scenario 2, highly anomalous situation. Terenghi 2019.

Sound example can be found at:
https://www.saralenzi.com/design-actions-material PSW: SonificationDesign2021

Scenario 3 – Repetition

In a regular situation, the five sounds of Scenario 3 – Repetition, five short impulse sounds based on concrete sound material, would play in synchronous at 10:00AM with the same duration of a few milliseconds.

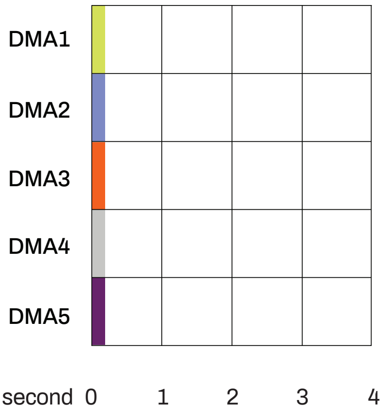


Fig. 40. Scenario 3, regular situation. Terenghi 2019.

Sound example can be found at:
https://www.saralenzi.com/design-actions-material PSW: SonificationDesign2021

At 11:00AM, at the appearance of some anomalous behavior in DMA 2 and DMA3, the corresponding sounds will loop within a maximum time interval that we capped at five seconds. The frequency of the loop would be faster or slower depending on the reconstruction error associated to each District.

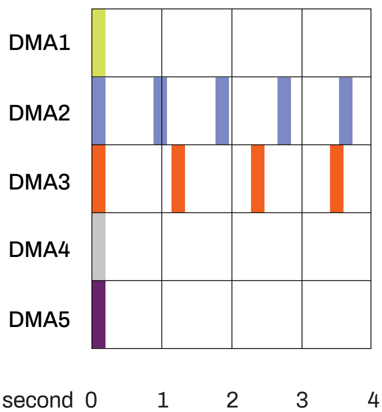


Fig. 41. Scenario 3, anomalous situation. Terenghi 2019.

Sound example can be found at:
https://www.saralenzi.com/design-actions-material PSW:SonificationDesign2021

At 12:00PM, as the attack progressively affects the network, all the Districts' sounds are looping with difference frequencies within the five seconds' duration of the sonification:

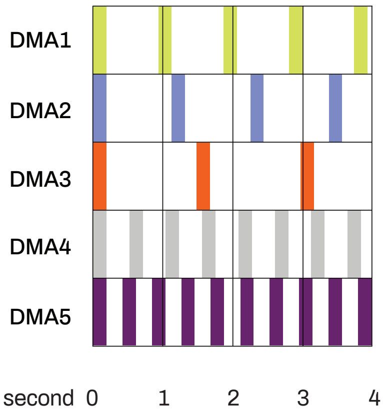


Fig. 42. Scenario 3, highly anomalous situation. Terenghi 2019.

Sound example can be found at:
https://www.saralenzi.com/design-actions-material PSW: SonificationDesign2021

Scenario 4 – Pitch

In a regular situation, the five sounds of Scenario 4 – Pitch, would play one after the other, starting from DMA1 to DMA5, from 10:00AM to 10:05AM. All districts will be played at the same pitch.

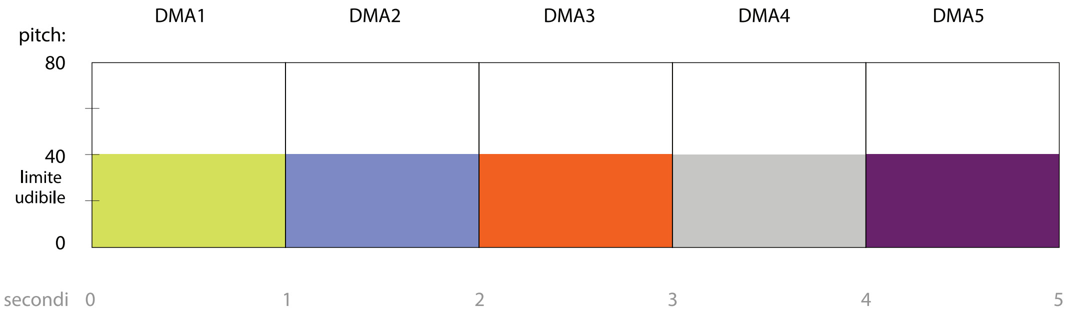


Fig. 43. Scenario 4, regular situation. Terenghi 2019.

Sound example can be found at:
https://www.saralenzi.com/design-actions-material PSW: SonificationDesign2021

At 11:00AM, at the appearance of some anomalous behavior in DMA 2 and DMA3, the pitch of the corresponding sounds will change, increasing proportionally to the value of the reconstruction error. In the example, the pitch for DMA2 is higher than DMA3.

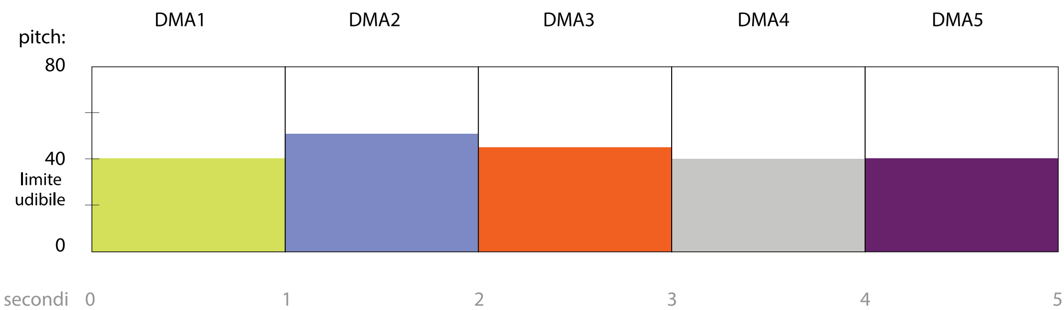


Fig. 44. Scenario4, anomalous situation. Terenghi 2019.

Sound example can be found at:
https://www.saralenzi.com/design-actions-material PSW: SonificationDesign2021

At 12:00PM, as the attack progressively affects the network, the pitch for all Districts has changed from the original one, played at 10AM.

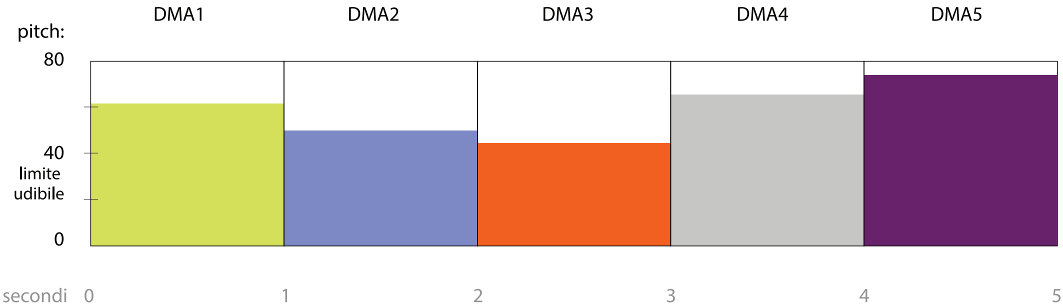


Fig. 45. Scenario 4, highly anomalous situation. Terenghi 2019.

Sound example can be found at:
https://www.saralenzi.com/design-actions-material PSW: SonificationDesign2021

In all scenarios, the sonification is repeated three times at a 10 second interval. We imagined that the first instance would attract the attention of the operator, the second would convey information and the third would be used to confirm the interpretation.

The four scenarios were shared with the team during a second feedback session. While the feedback was extremely positive on the overall change of strategy, some issues were identified mainly in scenario 1 – delay and scenario 4 – pitch. In scenario 1 the experts highlighted that the progressive loss of a reference point could interfere with the capability to rate the amount of anomaly in the districts. Moreover, in the absence of a signal identifying 'time zero', a listener could overlook the very start of the sonification. As a consequence, he could interpret the first delayed sound as the start of a regular sonification. As shown in Figure 36, in fact, when all the five districts are anomalous there is no sound played at the beginning of the sonification: the operator could interpret DMA3 as the first regular district. Additionally, he would have difficulties in comparing the relative amount of anomaly for each district in the absence of a starting reference point.

On the other hand, in scenario 4 the fact that the five districts are always played in sequence adds another task for the listener. Not only he has to remember the pitch associated to each district, but also its order.

We therefore decided to discard these prototypes and to move forward only with scenario 2 – Duration and scenario 3 – Repetition in their two versions: version A, which uses tuned sounds of instrumental origin and version B, which uses non – tuned sounds such as noise, concrete sounds, non-musical synth sounds and so on.

Prototype implementation- Scenarios 2 and 3 in version A and B were implemented by Terenghi (cit.) using the open-source Python script MIDI Time 1.1.3[25] which allows data values to be used to determine parameters such as the volume, the duration and the order of a sound sequence. The MIDI file would be read by the audio production software Ableton Live! 10[26] which would play the corresponding sounds, respectively:

25. https://pypi.org/project/miditime/

26. https://www.ableton.com/en/shop/live/

- Scenario 2A: chimes, from Ableton Live Sample Library.

- Scenario 2B: glass filtered noise with reverb, Ableton Live Instruments.
- Scenario 3A: piano little one, Ableton Live Instruments.
- Scenario 3B: samples from Ableton Live Drum Kits.

Experimental design

We then turned to investigate possible methods for the experimental phase. To recap, the prototypes had been designed with the following assumptions in mind:

- Through the sonification, the operator would be able to understand when the network is under attack i.e., when it presents an anomalous behavior.
- The sonification would also help the operator efficiently discriminate between different levels of anomaly, therefore allowing better decisions to be made on how to intervene.
- The operator would be able to locate the area of the network (specifically, the district) in which the anomaly is taking place.

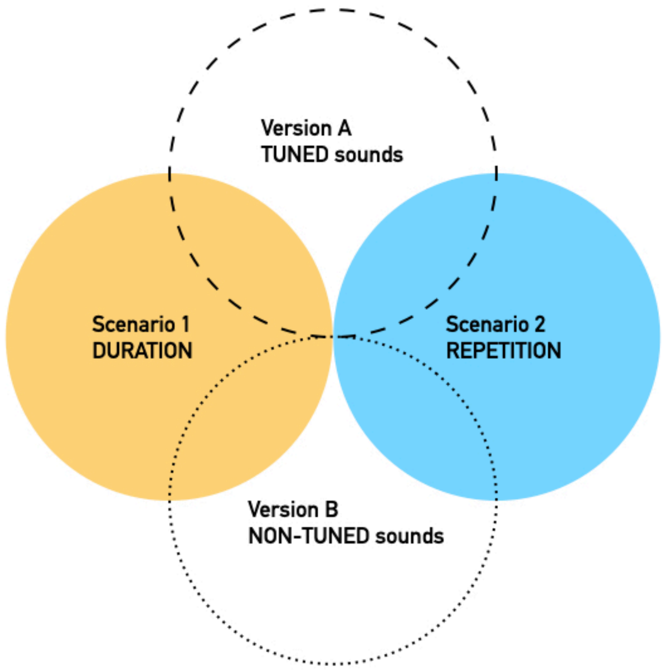


Figure 46. Combination of scenarios and versions evaluated in the experimental phase.

We rearranged the scenario names as follows (See Figure 46):

1. Scenario 1A: duration with tuned sounds
2. Scenario 1B: duration with non-tuned sounds
3. Scenario 2A: repetition with tuned sounds
4. Scenario 2B: repetition with non-tuned sounds

The goal of the experimental phase was to assess the performance of domain experts in identifying cyber-attacks, evaluate the gravity of the anomaly and locate them in the corresponding district, and then compare results across the four different scenarios. At the same time, we aimed to gather as much feedback as possible on the different prototypes in order to choose one for further development. Additionally, we were interested in how users would relate to sonification (a novel means of data representation for most domain experts), how they would integrate it within their daily routine and how they though it could be adopted in a real-world scenario. We therefore needed to design an experimental protocol which involved both qualitative and quantitative methods.

Qualitative Research- As the reader might recall, one of the research questions (RQ2) aims to explore and possibly define an experimental protocol for sonification projects. In Chapter 1, we claimed that the lack of a standardized experimental procedure and accepted experimental protocols are among the reasons preventing sonification from being recognized as a legitimate means for data representation. We therefore approached the qualitative part of the experimental phase with an open mind, in search of interdisciplinary sources of inspiration. For several reasons, among which, my personal doctoral research within a design department, a joint doctoral course led by TU Delft and dedicated to Research through Design (RtD) methods, as well as existing literature on research methods involving prototypes and sound design, we initially focused our attention on 'Technology Probes' (Hutchinson, Mackay, Westerlund 2003) and 'Design Probes' (Hogan and Hornecker 2016). Gaver and colleagues, who first introduced the concept of 'Cultural Probes' at the end of the 1990s, describe probes as a method used to "inspire developments in a design process" and to "provoke inspirational responses" (Gaver, Dunne, and Pacenti 1999, page 22). The fact that Gaver is a well-known scholar in the area of sound studies and contributed to the field of auditory display with seminal work (for example on auditory icons, see Chapter 2) seemed promising. Research based on probes is an "approach that values uncertainty, play, exploration, and subjective interpretation" as ways of dealing with the limits of knowledge (Gaver et al. 2004, p.53). In approaching design research for the first time I wanted to embrace a spirit that would preserve what the essence of design is for me: experimentation in practice, continuous adaptation to shifting problems, the effort of making implicit decisions explicit based on partially subjective considerations. Even more so when dealing with data sonification and therefore, to some extent, with the "subjective interpretation of objective, numerical values" (from the interviews with sonification experts, see Chapter 2).

Usually, probes are introduced at a very early stage of the conceptual development in order to gain fresh, unstructured insights that could even radically change the course of the prototyping. In the words of Hutchinson et al. (2003, p.1), who borrowed the concept of cultural probes and adapted it to the context of Human Computer Interaction's research on new technologies, probes satisfy three main goals: "The social science goal of understanding the needs and desires of users in a real-world setting, the engineering goal of field- testing the technology, and the design goal of inspiring users and researchers to think about new technologies." Again, we believed that these goals matched our objectives. The reader will recall that we did not have only one prototype to test, but four (or rather two prototypes in two versions). In this sense, our setting was close to the one described by Hogan and Hornecker (cit.) when introducing the concept of design probes: "I also consider it to be close in intent to Technology Probes however, instead of studying the use of one artefact (which is the procedure followed with technology probes), I create multiple artefacts that possess similar design features but differ in one aspect. This allows researchers to focus the evaluation precisely on this design feature - in my case this was representational modality." (p.5). In our case, this was the mapping strategy and the choice of sound material.

Once we had identified the overarching framework for the qualitative part of the research, we listed several areas to explore during the experiment. In particular:

- User Experience: technical issues, equipment, testing environment, working activity during testing. Feedback on the potentiality of the sonification in the real world.
- Sound Design: feedback on the choice of sounds, the structure of the different scenarios, the overall listening experience (what worked, what didn't work).
- Post-Quantitative Testing: we would share the results with each tester and comment on these during the interview.

Quantitative Research- In this phase, the goal of the quantitative testing was mainly to assess and compare the ease of use of each prototype. In particular, we wanted to validate:

- The use of sound to communicate relevant information on the status of the network.
- The capability of the operator to correctly understand the meaning of the information conveyed by sound.
- The most promising prototype among the four options.

The goal of testing was not, at this stage:

- To obtain statistically relevant information on the accuracy of the sonification, where accuracy is defined as the capacity to effectively reach one's goal.
- To obtain statistically relevant information on the efficiency of the sonification, where efficiency is defined as the capacity to reach one's goal using the minimum resources.

These two (very relevant) aspects were set aside for a future possible round of testing with a much wider user base.

We designed the test around a preliminary question, three main questions and a fourth optional question. The testers would answer the questions just after hearing the sonification every hour.

x) Did you hear the sound? Yes/No

i) If not, why? (Were you busy with something else? Where you too far from the computer? Was there too much noise around you?)

[If the answer is “No” the questionnaire stops]

This preliminary question was meant to exclude technical problems which would prevent the tester from completing the remaining questions and help us better understand the flow of the sonification during a typical working day.

1) How would you describe the status of the system? Anomalous / Regular

[If the answer is “Regular” the questionnaire stops]

The first question aimed at evaluating whether the user could identify anomalies in the network, in that specific moment.

2) How would you rate the anomaly level on a scale 1 - 5 where 1 = least serious and 5 = most serious?

The second question aimed at evaluating whether the listener could evaluate the gravity of the anomaly through sound. As mentioned, no pre-determined level of anomaly was embedded in the sonification. Rather, we hoped that users would, over time, attribute an anomaly index based on their knowledge of the system. The spirit of this question was to validate our hypothesis by evaluating the consistency of the rating through different users.

3) How many Districts present anomalies?

* One District
* Two Districts
* Three Districts
* Four Districts
* All Districts

The third question would be used to compare the efficacy of the different scenarios against one of the goals of the sonification: to facilitate locating the anomaly on the network geography.

4) Which district presents anomalies and what is the level of the anomaly? [select the district/s which presented an anomaly and specify the anomaly level].

* D2 Anomaly level _____ [1 to 5]

* D1 Anomaly level ______ [1 to 5]
* D3 Anomaly level ______ [1 to 5]
* D4 Anomaly level ______ [1 to 5]
* D5 Anomaly level ______ [1 to 5]

The fourth question, which we originally considered optional, aimed to refine the evaluation of the capability of the user to locate the origin of the anomaly within the network, a preliminary step for a more efficient retrieval of additional information from the visual map of the SCADA. We also asked the testers to attribute an individual level of anomaly to each district although we were aware this was going to be a difficult task. However, since all the testers answered this question, results will be fully reported in the analysis but will be kept separate to maintain the original spirit of the 'optional' question.

Experimental Setting- Following the design probes approach, we decided to deploy the experiment "in a real living and working environment" (Hogan and Hornecker, cit.). The experiment included three phases:

- A preliminary questionnaire, to assess the competence of each user in the fields of water management and cyber-threats and their familiarity with music and/or sound production.
- A task-oriented test to be completed during the experiment, to assess the performance of the users in understanding the status of the network, rating the level of anomaly on a 1 to 5 scale and identifying how many districts (and possibly which district) were anomalous.
- A final, semi-structured interview where we aimed at gauging feedback on the overall experience, design and technological issues and gathering suggestions for future developments of the prototypes.

Six experts from the field of water infrastructure management and cybersecurity (three men and three women, from different countries and cultural backgrounds) were selected. Results of the preliminary questionnaire, where experts were asked to self-assess on a Likert scale (from 1 to 5) their competence in water management and cybersecurity and their familiarity with music and more in general, with sound, are shown in Figure 47.

Domain expertise self-assessment

TN: tester ID
area of expertise
age

Work sector:
▽ academia
○ industry

Likert Scale:
0 No specific expertise
5 High expertise

CS: cyber security
WM: water management
M: music

● specific competence
○ passion

Gender:
▼ male
▼ female

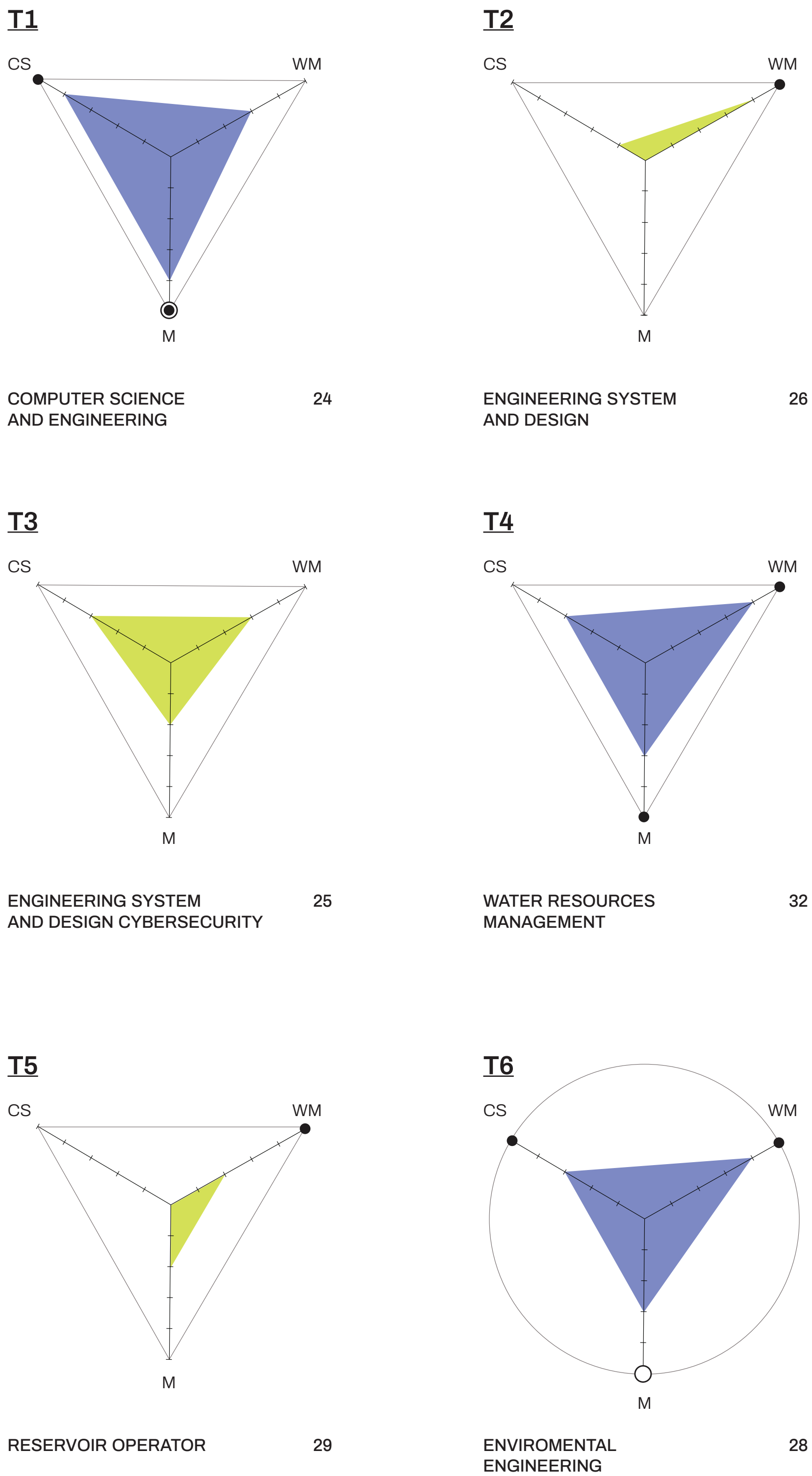


Fig.47. Results of the preliminary questionnaire. Adapted from Terenghi cit.

On testing days, scenarios 1 (A and B) and 2 (A and B) were made available to the testers via real-time streaming on a purposely designed webpage. Our users accessed the web application at the beginning of their working shift. The sonification played automatically every hour over eight consecutive hours of a normal working day. Every hour, the sonification played for a maximum of 10 seconds. It was then repeated three times, to make up for distractions or to allow the operator to clarify potential doubts or confirm an interpretation. Testers would listen to the sonification through their favorite sound system, either headphones or speakers. After hearing the sonification every hour, they were asked to fill in an offline spreadsheet that we provided to answer the questions. The use of real-time streaming instead of pre-recorded sounds uploaded on a web player ensured testers could not listen to the sonification before or after the planned time or listen to it more than once.

In line with the probes approach, we decided to provide users with a general introduction to data sonification and to the specific use case, but did not require them to go through a training phase or provide the testers with a demo of the scenarios. This way, we hoped to inspire users to find their own way in the relationship with the prototypes in order to collect as much unexpected feedback as possible. Only the keys along with the individual sounds for each district were provided. All the material was uploaded on a purposely designed web page and is accessible to the reader at the following link: https://ginevraterenghi.github.io/sonifying-cyber-attacks/p3.html.

Having four different scenarios to test over a few days, we considered the risk that the testers would learn from the first sonification and that this would inevitably improve their performance with the following ones. We therefore decided to schedule the testing of each scenario over a period of two weeks to have a few days' interval between each test and randomize the order of the prototypes, according to the following schedule:

1. Monday, Week 1: Scenario 1A
2. Thursday, Week 1: Scenario 2B
3. Tuesday, Week 2: Scenario 1B
4. Friday, Week 2: Scenario 2B

Additionally, we chose two different datasets with different anomalies, one for scenarios 1 and one for scenarios 2.

Risks and limitations- The testers familiarizing themselves with the first prototypes and performing better in the testing of the last one was not the only possible bias we considered. In general, design research methods such as probes tend to collect subjective considerations which make it difficult "to interpret, let alone analyze" (Gaver, Dunne, Pacenti cit.). Additionally, according to 'The Design Exchange'[27] this method is highly dependent on the motivation of participants as they are supposed to reflect upon and be inspired by the probes, "That's why it can be hard to motivate participants, and so there is a risk of few returns; it can be hard to get them to send the probes back." (The Design Exchange 2020). In our approach, by not providing training or demos we hoped to inspire participants to experiment and find their own way. The choice of allowing the participants to freely explore the prototype without having them undertake a preliminary and potentially tedious training session was taken also in an attempt to motivate them to develop their own way of relating to sound. In the following paragraphs we will show how each participant tended to relate the sonification to pre-existing experiences, an attitude which we believe increased their engagement during the whole experiment. In terms of the quantitative testing, the major limitation clearly lies in the number of participants (six users) which makes the results far from being statistically relevant. Nonetheless, we believe that the quantitative, performance-related tests together with the semi-structured, qualitative material collected during the interviews allowed us to build a solid case for an integrated evaluation of the four prototypes, especially since the testers were all domain experts.

27.
The Design Exchange is an online resource on design research methods. It can be found at: https://www.thedesignexchange.org/.

Results

Quantitative analysis- The analysis of the quantitative testing focused at first on the three main questions and later on the fourth optional question. Figure 48 shows the four different layers of analysis in hierarchical order. Priority was given to identification of anomalies in the system, followed by the capacity to attribute a level to the anomaly, distinguishing between serious anomalies requiring immediate intervention from less serious ones. The third layer analyzed the identification of the number of anomalous districts. The least priority was given to the capacity of the listener to identify anomalous districts and attribute a level of anomaly to each of them.

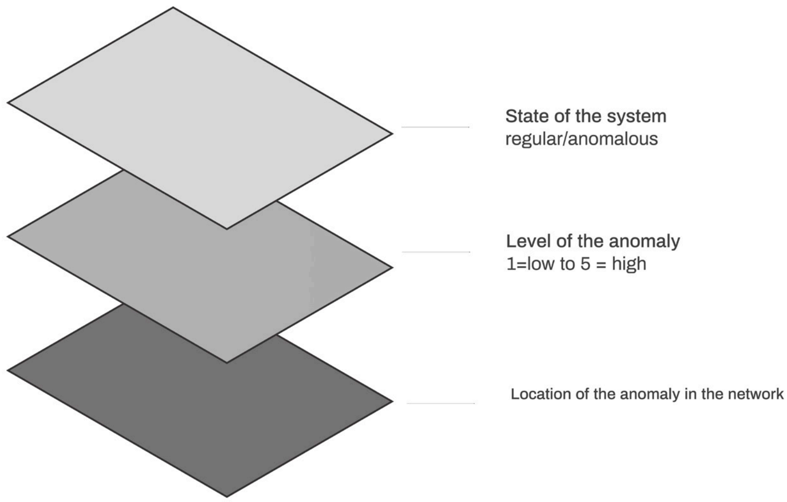


Figure 48. Layers of quantitative analysis. Terenghi 2019.

To follow the analysis more clearly, the reader can refer to Figure 49 which reports the correct reading of the network's anomalies over the four days of testing. Darker colors correspond to a higher level of anomaly while the color white indicates that the system is behaving regularly. The network always presents a certain level of anomaly except for the first and the last hour of testing. In both scenarios, anomalies are concentrated in the central hours of the day with scenarios 1 showing a higher severity of the attack (see hour 3, Scenarios 1) while scenarios 2 are affected by a lower anomaly level. In both groups the attack affects various districts in the network, with scenarios 2 showing a more intermittent behavior of the anomaly, with the number of affected districts increasing, decreasing and then increasing again (hours 4, 5 and 6 of layer 3).

Results vary along the three layers and four scenarios. The identification of the anomalous status of the system didn't seem to pose any issue in any of the scenarios, but for the scenarios 1, which are based on duration, three of the testers misinterpreted regular data as anomalous, at the very beginning and at the very end of test, as shown in Figure 50.

How to read

☐ anomalous data

hn. hour to which data refers to

regular
level 1/min
level 2
level 3/mid
level 4
level 5/max

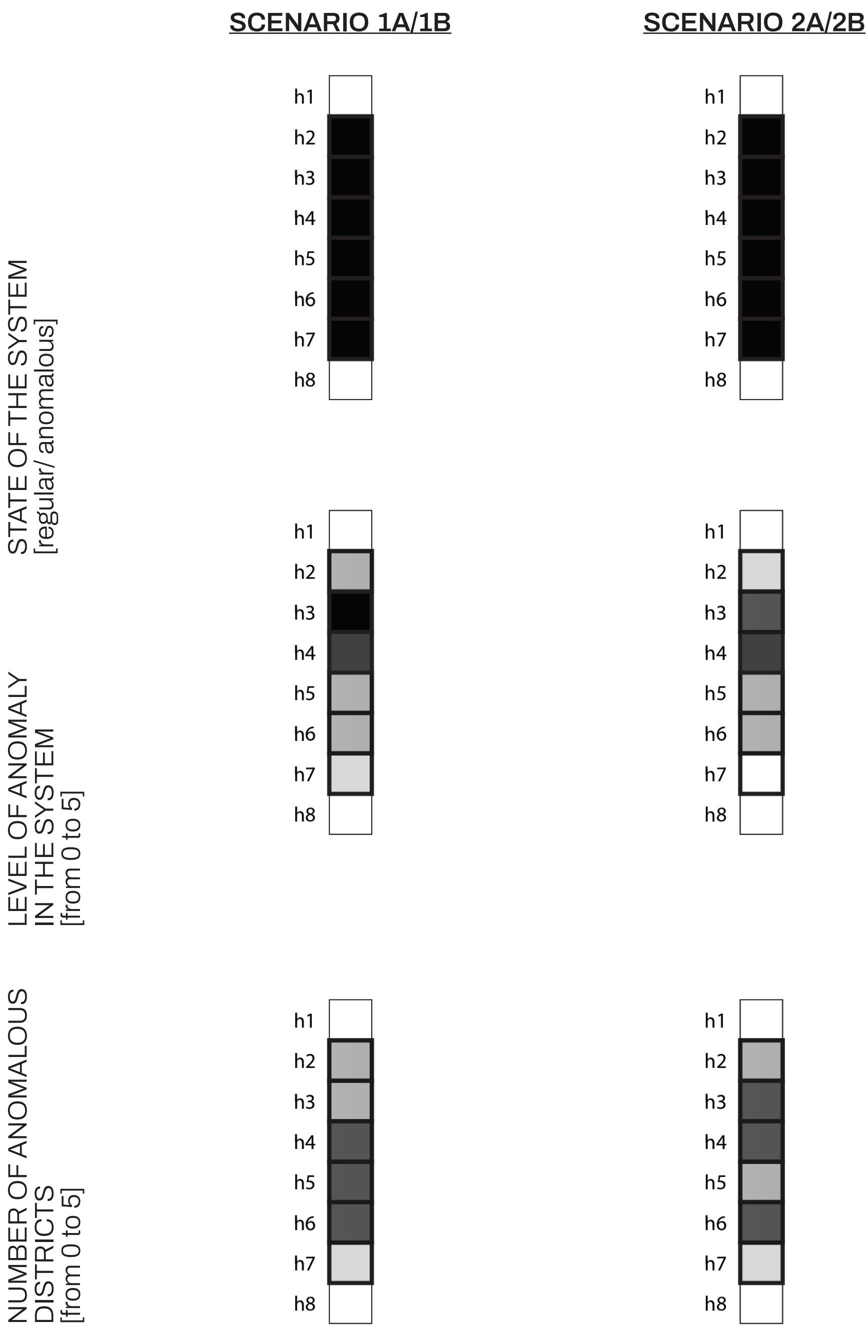


Figure 49. Correct answers for each layer. Terenghi 2019.

How to read

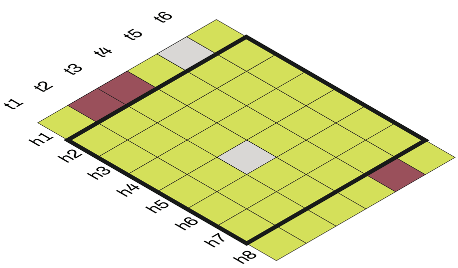
anomalous data
tn. tester ID
hn. hour

correct answer
wrong answer

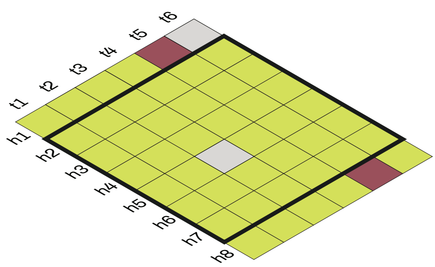
no answer
non-reliable answer

STATE OF THE NETWORK
[regular/ anomalous]

SCENARIO 1A

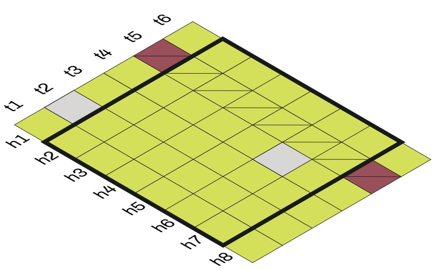


SCENARIO 1B



STATE OF THE NETWORK
[regular/ anomalous]

SCENARIO 2A



SCENARIO 2B

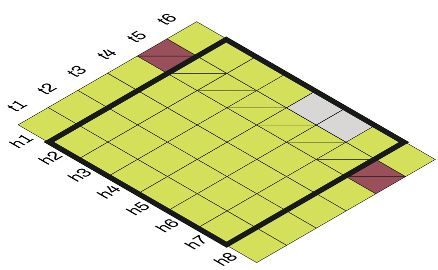


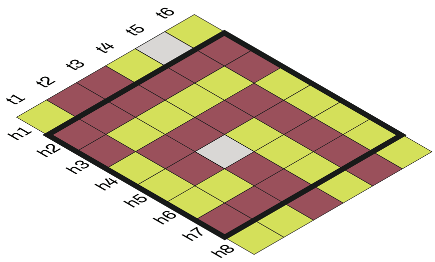
Figure 50. Results, layer 1. Adapted from Terenghi 2019.

How to read

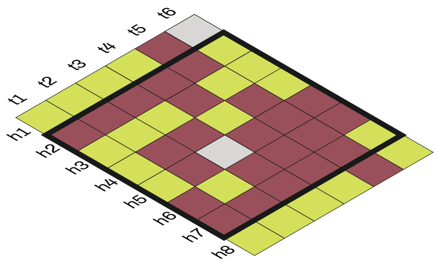
- anomalous data
- tn. tester ID
- hn. hour
- correct answer
- wrong answer
- no answer
- non-reliable answer

LEVEL OF ANOMALY IN THE NETWORK
[from 0 to 5]

SCENARIO 1A

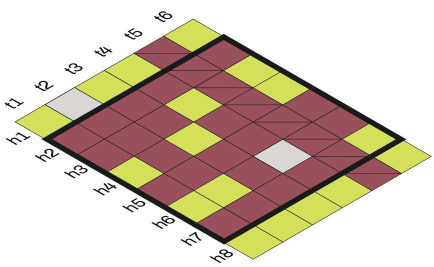


SCENARIO 1B



LEVEL OF ANOMALY IN THE NETWORK
[from 0 to 5]

SCENARIO 2A



SCENARIO 2B

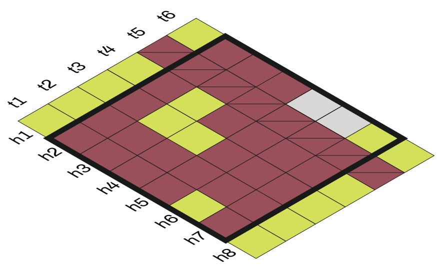


Figure 51. Results, layer 2. Adapted from Terenghi 2019.

Two of them (testers 2 and 3) realized the mistake shortly after completing the spreadsheet and corrected it on a separate note. For scenarios 2, no errors were recorded: all testers clearly distinguished between regular and anomalous behavior, with the exception of tester 5 who later realized she had misunderstood mapping rules for scenario 2 and asked for her answers to be withdrawn from the analysis. While keeping in mind the extremely limited statistical significance of a test conducted with six users, we can appreciate that at this layer, the type of sound material does not seem to influence the listener: the testers' performance is consistent, both with tuned and non-tuned sounds.

Results are more complicated when we have a closer look at the second and third layer i.e., when we ask testers to rate the anomaly level of the network and identify the number of anomalous districts. At layer 2, as Figure 51 shows, the performance is more irregular. The reader should note that, as I previously mentioned, the dataset does not contain an indication of the level of anomaly. Such level was obtained retroactively while analyzing the testers' answers to check their performance. It was a choice by the research team to ask the listener to attribute an anomaly level in order to assess how much information the operator would gather from the sonification. This assessment would constitute a first pool of data for further evaluating the efficiency of the sonification both in facilitating operators taking action in the event of a real emergency and their capacity to distinguish a false positive.

Looking at Fig.51, scenarios 1 (based on duration) seem to perform better. During the follow-up interviews some testers lamented the lack of a reference scale against which to compare the level of anomaly in scenarios 2 ("For the repetition's scenario, there was no relative scale for comparison"). In scenarios 1, the maximum duration of a sound was 10 seconds (the duration at which the whole sonification was capped), thus representing the maximum level of anomaly. In scenarios 2, it is the frequency of repetitions over 10 seconds which increases as the anomaly level increases. This increase is perceived by the listener as a higher speed of certain sounds compared to others, but it is very difficult, or even impossible, for the listener to judge beforehand what the maximum number of repetitions is over 10 seconds. Perhaps as an unconscious emotional reaction to increased speed, the anomaly level for scenarios 2, both A (with tuned sounds) and B (with non-tuned sounds) has been overestimated rather than underestimated. Figure 52 shows the direction of the error towards an over- or underestimation of the anomaly level on a scale of 1 to 5: in grades of purple we can see how the error corresponds to an underestimation of the gravity whereas in a gradient of orange, how much the anomaly has been overestimated.

If answers for scenarios 1 seem to be more balanced between under- and overestimation, scenarios 2 show a clear overestimation of the anomaly level, with an additional increase for scenario 2B, where the sound type was non-tuned, concrete sounds mainly sourced from various types of mechanical noise. Interestingly, despite the poor results, most of the testers seemed to be more engaged emotionally by scenario 2B, which they judged “fun”, “easy to remember”, “easier to understand” and “playful”.

Results are reversed for the third layer, in which the listener was asked to determine the number of anomalous districts. As shown in Figure 53, the error rate for scenarios 1 is clearly higher than for scenarios 2.

Figure 54 also tells us than in both cases the number of districts affected by the anomaly was mainly underestimated rather than overestimated. In the follow – up interview, testers reported that they found it more difficult to distinguish individual sounds in scenarios 1 when more than one district was anomalous.

With more than one district presenting anomalies, the duration of more than one sound increases, thus creating superimpositions that make it more difficult to isolate an individual element. This effect was less evident in scenarios 2, where sounds repeat over time maintaining their original duration (one second), making it easier for the listener to distinguish each sound individually.

The test included a fourth, optional question which we considered particularly challenging. We asked the testers to identify the level of anomaly for each district, for each hour. We welcomed as greatly positive that all testers accepted the challenge and answered this fourth question. Figure 55 illustrates the results: for each of the eight hours of the test, it reports the answer of each tester for each district and the indication of correct (in green) or wrong (in brown) answer.

QUANTITATIVE ANALYSIS:

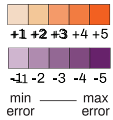
relevance and error’s direction

How to read

sonification contains anomalies
tn. tester number
hn. hour number

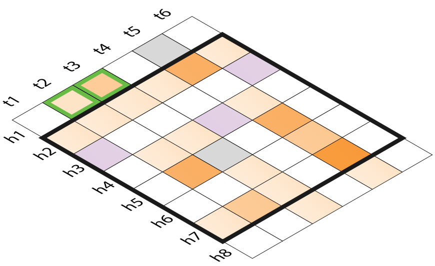
correct answer
no answer
not reliable
tester made a mistake which was later amended

ERROR RELEVANCE
color grading represents the error’s direction
i.e., whether the tester over or understimated the gravity of the anomaly

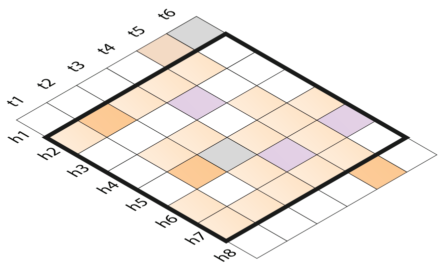


LEVEL OF ANOMALY IN THE NETWORK
[from 0 to 5]

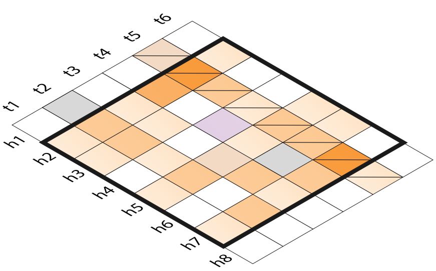
SCENARIO 1A



SCENARIO 1B



SCENARIO 2A



SCENARIO 2B

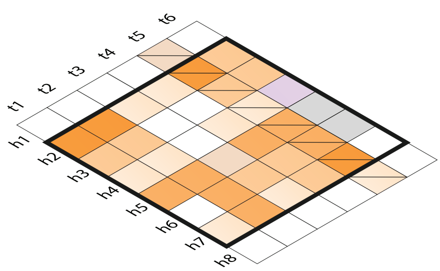


Figure 52. Direction of the error for layer 2. Adapted from Terenghi 2019.

How to read

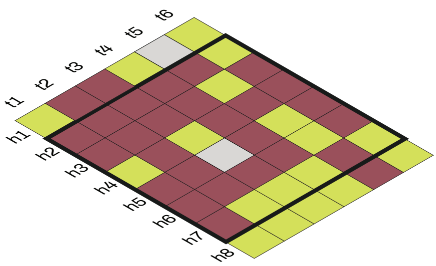
anomalous data
tn. tester ID
hn. hour

correct answer
wrong answer

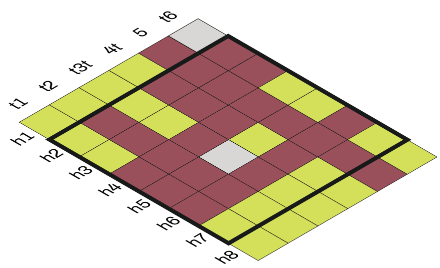
no answer
non-reliable answer

NUMBER OF ANOMALOUS DISTRICTS
[from 0 to 5]

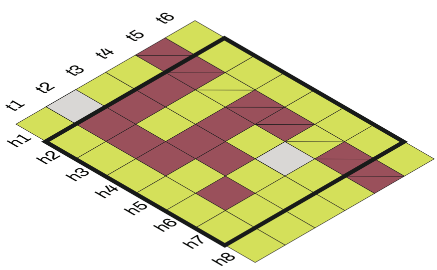
SCENARIO 1A



SCENARIO 1B



SCENARIO 2A



SCENARIO 2B

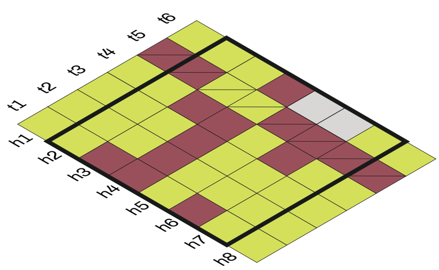


Figure 53. Results, layer 3. Adapted from Terenghi 2019.

QUANTITATIVE ANALYSIS:

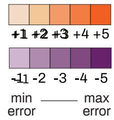
relevance and error's direction

How to read

sonification contains anomalies
tn. tester number
hn. hour number

correct answer
no answer
not reliable
tester made a mistake which was later amended

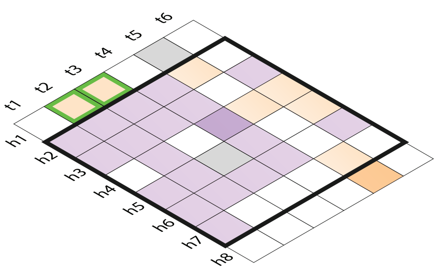
ERROR RELEVANCE
color grading represents the error's direction i.e., whether the tester over or understimated the gravity of the anomaly



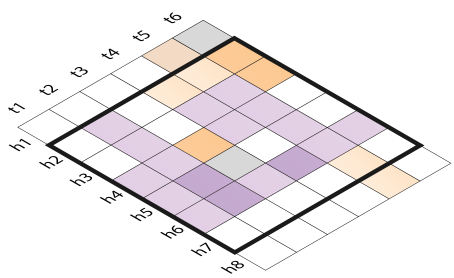
NUMBER OF ANOMALOUS DISTRICTS

[from 0 to 5]

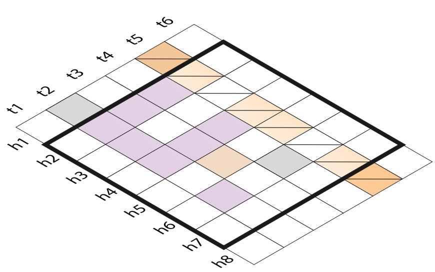
SCENARIO 1A



SCENARIO 1B



SCENARIO 2A



SCENARIO 2B

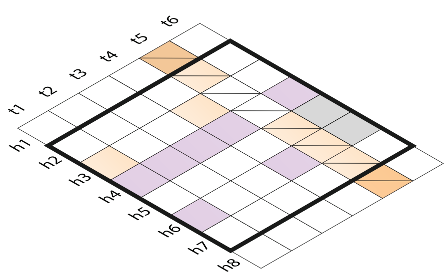


Figure 54. Direction of the error for layer 3. Adapted from Terenghi 2019.

Anomalous districts at every hour are highlighted by a thicker border.

As expected, the highest rate of error occurs in the central hours, where the network is under attack. Answers are consistent across the different testers and the different scenarios. At H2, scenarios based on tuned sounds (labeled with 'A') seem to perform slightly better than scenarios based on non-tuned sounds, whereas at H4 and H5, the answers for scenarios 2 (repetition) are slightly more correct.

Figure 56 shows that, again, answers for scenarios 1 tend to underestimate the level of anomaly whereas answers for scenarios 2 tend to overestimate it.

Needless to say, the limited sample base does not allow for conclusive results. In general, it came as no surprise that the error rate is higher when both the number of districts involved and the level of anomaly in each district are higher and therefore, the situation in the network is more confused as the attack is causing more damage. I remind the reader that we designed the sonification as a complementary tool to the visualization of data in the SCADA system. We hypothesize that in a real-world situation, the operator would react to the perception of an increasingly serious situation, where more districts are affected, by referring to the analytical information provided by the SCADA system. In other words, we believe that, in this context, the sonification need not to be analytically precise but rather 'good enough' to trigger the appropriate reaction in the operator rather than conveying specific information on each district. I will further elaborate on this point in the final part of this work, Chapter 6.

A separate reflection should be dedicated to the perception of the anomaly level. We hypothesized that the operator would be able, over time, to identify nuances in the anomalous behavior and that this ability would help the operator identify possible patterns, make predictions on future alarms and possibly discriminate between true and false alarms issued by the algorithm. It was with this hypothesis in mind that we decided to assess the capacity of the user to discriminate between levels of anomaly. During the follow-up interviews several testers pointed at

How to read

▭ anomalous data
tn. tester ID
hn. hour
DN. DISTRICT

correct answer
wrong answer

no answer
non-reliable answer

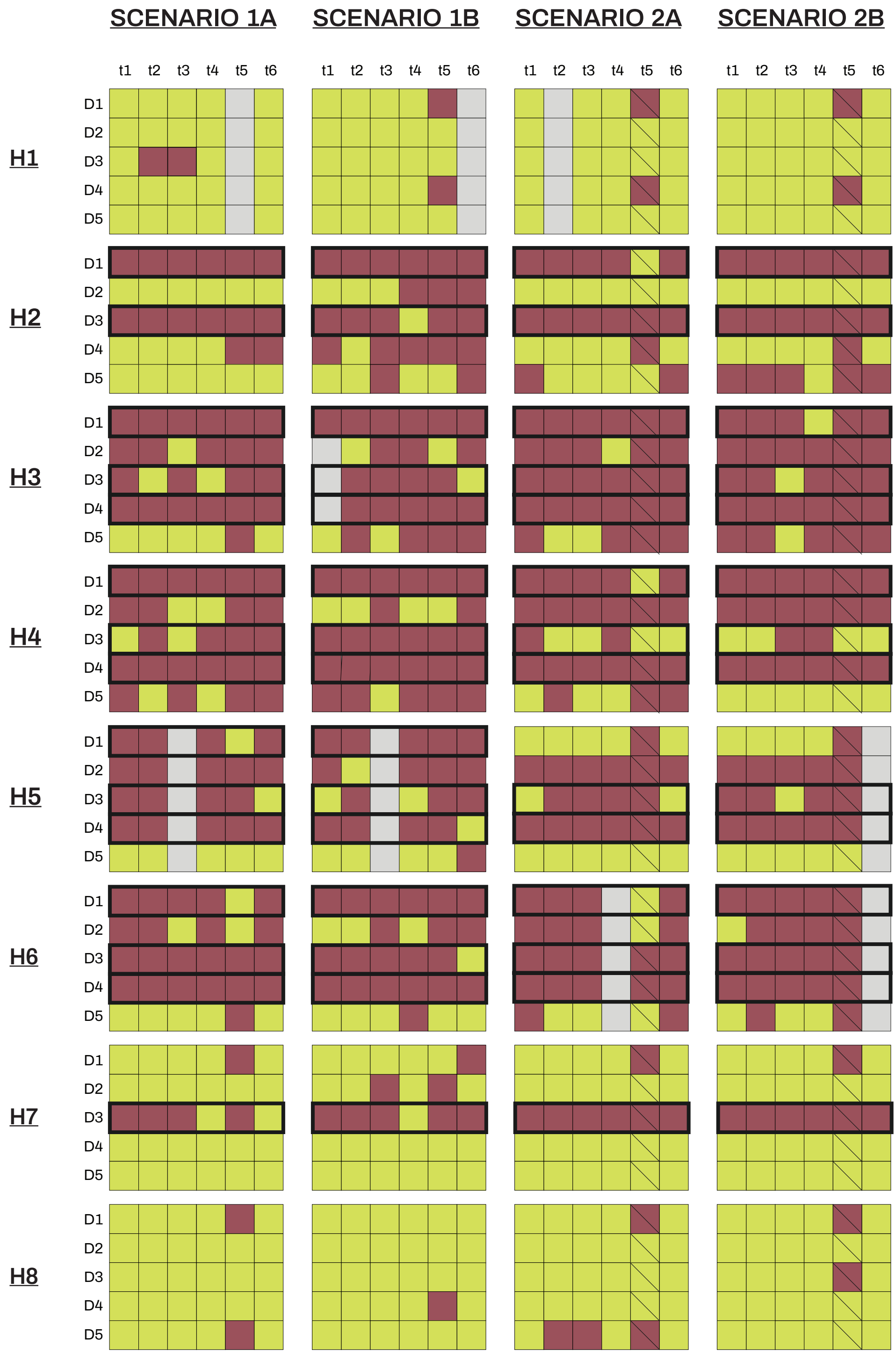


Figure 55. Results, layer 4. Adapted from Terenghi 2019.

QUANTITATIVE ANALYSIS:
relevance and error's direction

How to read

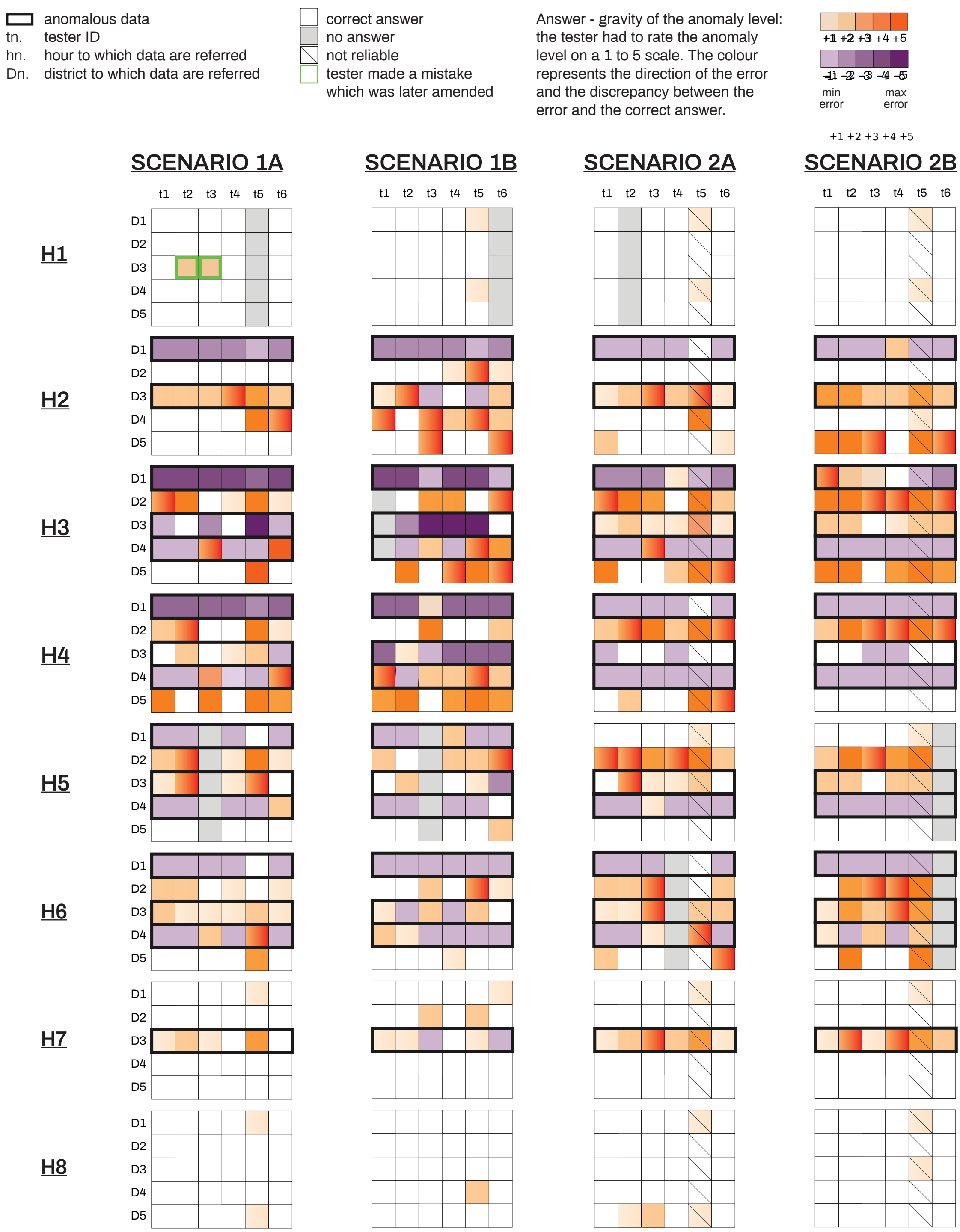


Figure 56. Direction of the error for layer 4. Adapted from Terenghi 2019.

the fact that, listening to the sonification and in the absence of an absolute reference scale, “the judgement of the severity of the anomaly is subjective” and as such, they found it difficult to scale it over a 5 – step scale, as we required. We understand this feedback as a confirmation of what we aimed at with the design of the sonification. In principle, we designed the sonification as a tool that helps the operator leverage his personal and subjective knowledge of the network. A knowledge accumulated over several years of experience supervising the functioning of a vital infrastructure in what is often the city where operators work and also live. Possibly, in their vision of the network, the everyday experience of the urban design of the city is superimposed with their mental map of the water network. While the results of the quantitative testing tell us that no scenario presents clear advantages and that several adjustments should be made for the design of a second iteration of the sonification, we welcome as positive the capability of the testers to recognize anomalies, recognize when more than one district is involved and understand when the anomaly level changes. We also positively welcomed the emergence of the idea of a subjective understanding of some of the aspects in the dataset as well as the idea of sonification as a means to represent ‘good enough’ information that can be later integrated with analytical knowledge provided by other – mainly visual – sources of information. We believe this is a promising result in the direction of using sound to facilitate human operators in taking better, more informed decisions without relying only on information provided by Artificial Intelligence.

Once we had completed the analysis of the quantitative results, we had high expectations regarding the insight the qualitative data we collected during the follow-up interviews would provide.

Qualitative analysis- The follow-up interviews were semi-structured around pre-determined areas which we wanted to investigate. In particular, we wanted to gather as much feedback as possible on the testers’ experience of usage during the whole experiment; on the sound design of the sonification; on the overall understanding of the sonifications, in particular in relation to the identification of the districts (a point that had raised some issues in the quantitative testing); on the choices we made in terms of the duration (ten seconds) and repetitions (three times) of the sonifications, their frequency (once every hour) and finally, on the potentialities of the application of the sonification to a real-world situation. Finally, we wanted to try to gather specific feedback and suggestions on each of the four scenarios and comment on the testers' results in the quantitative experiment.

Experience of usage

This area helped us understand whether the testers had encountered specific technical problems during the experiment, comment on them and find possible solutions. Figure 57 recaps the comments of all the participants clustered around three main themes: what was their main occupation during the experiment; what kind of sound diffusion system they used; how was the outside environment and other potential critical aspects.

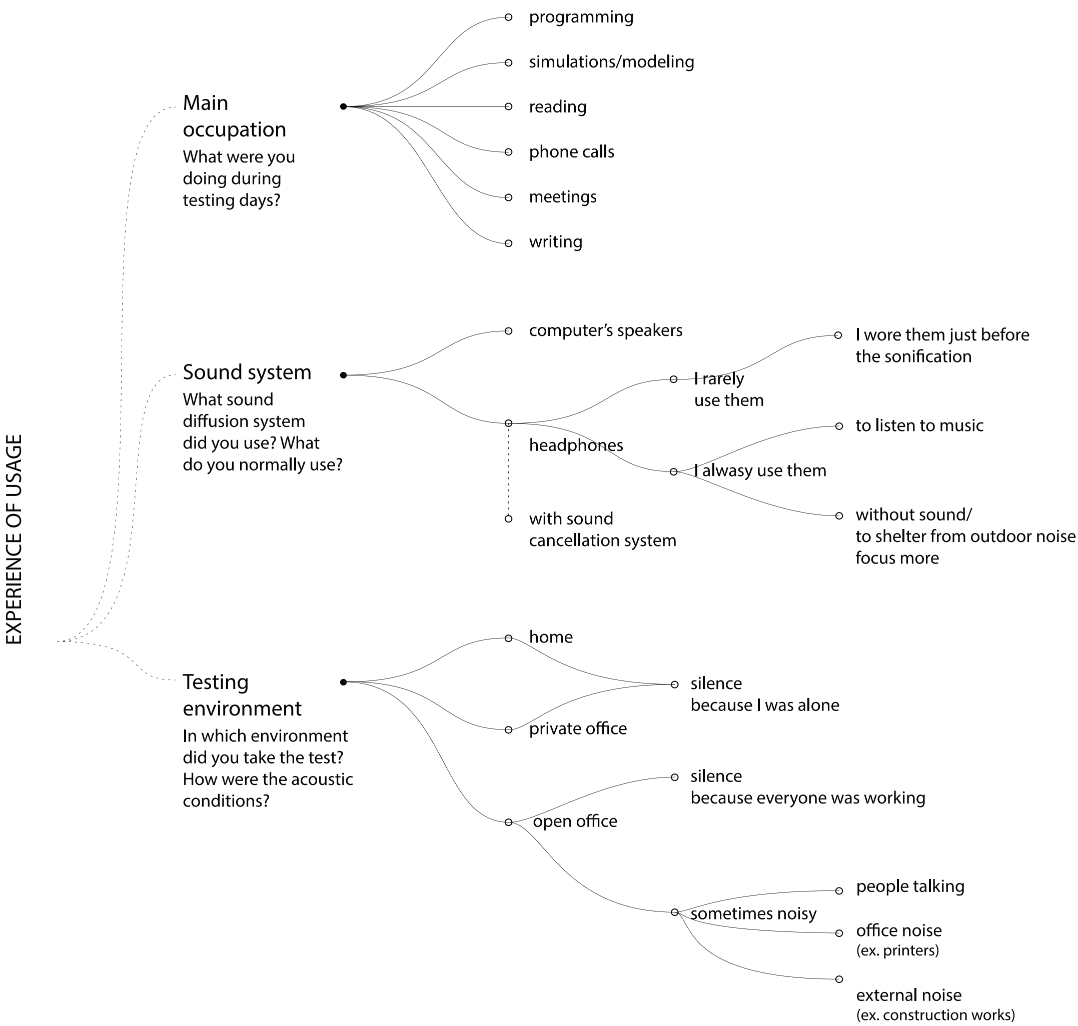


Figure 57. Results of qualitative research: User Experience. Adapted from Terenghi 2019.

The majority of the participants declared they use headphones at work, both with sound - to listen to their favorite music - and with no sound - to shelter from external distractions or noise. The vast majority of the participants (all except one) share their working space with colleagues in an open-plan office. One of the participants took one of the four experiments from home, due to changes in the office work schedule. The widespread use of headphones came as a surprise and represents a very important design constraint to be taken into account both for the design of future iterations of this prototype and for other sonifications based on a similar use case.The daily routine of the participants includes activities such as writing reports, doing research on historical data, meeting with colleagues and answering the occasional phone call. All the activities were consistent with the use case we had outlined at the beginning of the project and came as no surprise. One participant missed three sonifications (i.e., three hours out of 32 hours of the experiment) as she had to leave her desk for meetings and personal needs. This is something to take into account in the design of a real-world application and that could be solved by integrating the desktop sonification tool with a portable version (for example, for smartphones) or with a portable alert device. Another participant had to heavily reduce the volume of the sonification as an unexpected Skype call came through the same headphones as the sonification was playing. This is an aspect that could be easily solved with an automatic volume control which reacts to other incoming sounds by adjusting the sonification volume. Other technical issues were limited to a lack of synchronization between the computer clock and the sonification application clock, which prevented the sonification playing on time on the hour at the beginning of the experiment. The issue was easily solved by the iTrust team as the participants who had encountered the issue started the test again the following day.

In the following paragraphs, I share and comment a series of graphic representations in which the participants' thoughts that emerged during the interviews are clustered according to the main interview topics. Figure 58 shows the key to the representations: negative comments are indicated by a purple point, positive comments by a green point and neutral comments by a blue point. Suggestions are marked in grey. The size of the point associated with the key sentence indicates the number of participants who expressed the same thought with similar words.

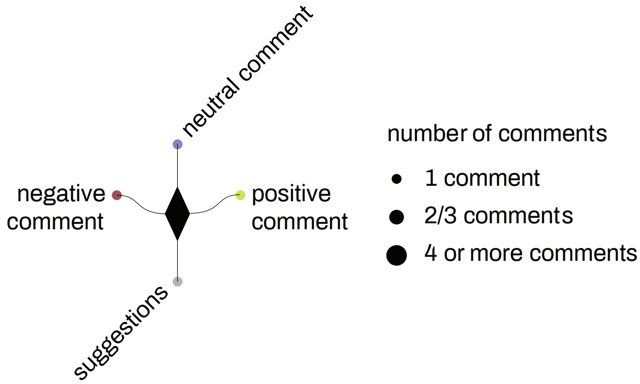


Figure 58. Key for the interpretation of qualitative analysis. Adapted from Terenghi 2019.

Sound Design

In the context of the overall objectives of this dissertation, the area of sound design is of particular relevance. We grouped around this topic comments and suggestions that describe how participants related to the use of sound in general and in particular to the specific types of sound we used. As shown in Figure 59, more than one participant admitted finding it difficult to associate sounds to a specific, individual meaning, such as the identification of a specific district. This echoes a finding from the set of interviews with authors of sonification, where more than one interviewee highlighted that sound can be used effectively to communicate messages on a phenomenon but things get complicated when we want to "represent hard values" (see Chapter 2).

A participant pointed out that there was a direct relationship between what he liked/disliked and aesthetic considerations ("I chose the preferred scenario mainly based on the aesthetics of the sound"). We will return to the relationship between feelings of like and dislike and aesthetic considerations in the following sections. More than one participant suggested using sound of different musical instruments to differentiate the districts, while one participant suggested adding a narrating voice to guide the user and reduce reaction time. Another participant suggested the sonification only play a sound in the event of an anomaly, a solution we consciously decided not to pursue during the design phase. In fact, the user would be unable to distinguish silence due to the absence of anomalous behavior from silence due to a technical failure of the audio equipment. In terms of emotional response to sound, a participant suggested we explore the usage of 'annoying' sounds to represent anomalies and 'pleasant' sounds to represent regular values, while another suggested we design the sonification as a shared monitoring device to be broadcast in the whole control room and in the whole water plant building.

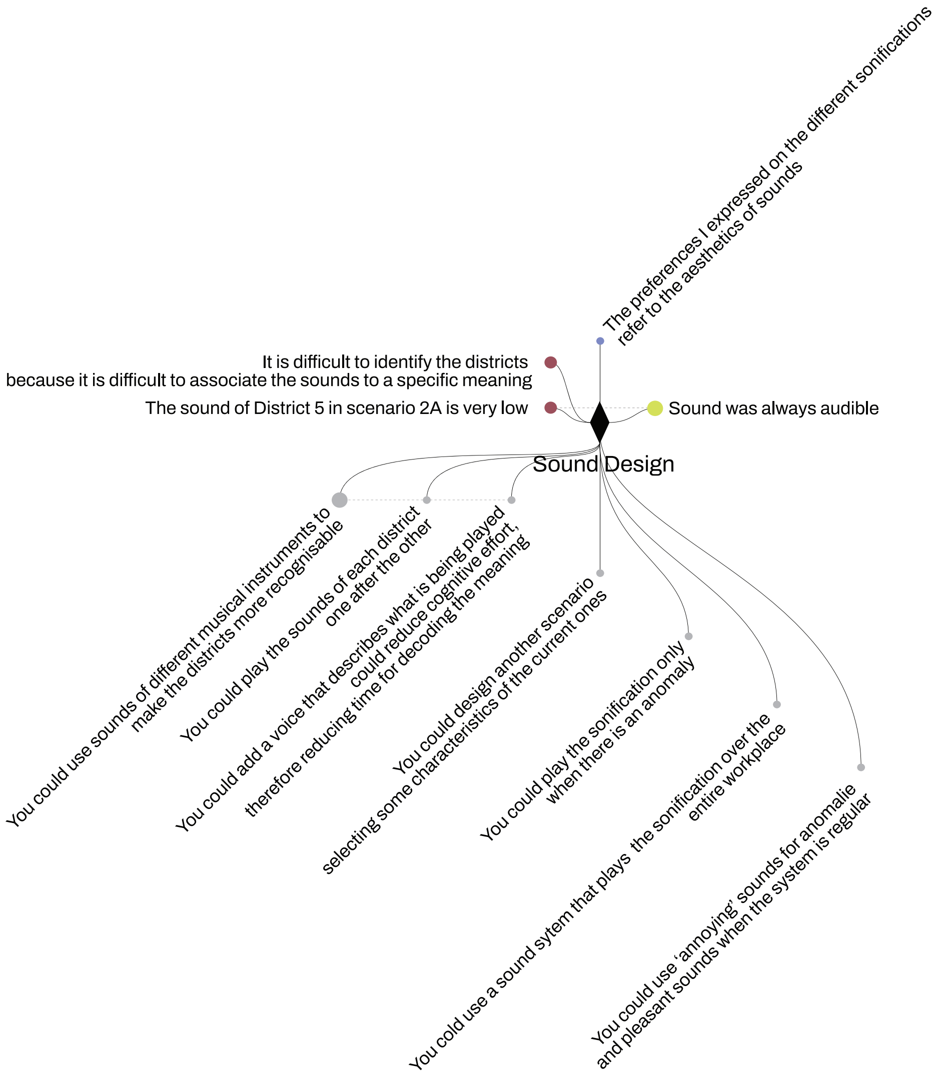


Figure 59. Results of qualitative research: sound design. Adapted from Terenghi 2019.

Overall understanding, district identification

The identification of the five districts presented some issues in the task-based quantitative testing. During the interviews, participants confirmed that difficulties increased when sounds of different affected districts tended to superimpose one another and that some scenarios were more difficult to grasp than others. Sometimes, they had to pause from their task to pay specific attention to the sonification in its three repetitions and only felt confident they had understood the message with the final repetition. As previously discussed, one of the participants raised the important issue of subjectivity in the interpretation of sound. Overall, all participants agreed that the identification of a district was more difficult than recognizing anomalous behavior in the overall network. On the positive side, which is identified by the color green in Figure 60, participants highlighted that the limited duration of the sonification for regular behavior (one second) made the system particularly efficient and useful for getting a quick overview while carrying out other tasks. Information conveyed by sound was generally reported as being easy to understand.

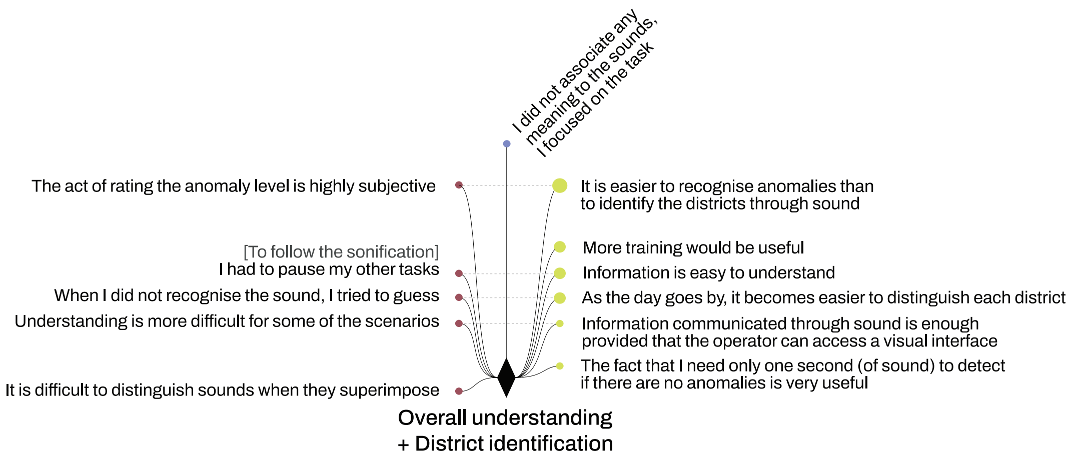


Figure 60. Results of qualitative research: district identification. Adapted from Terenghi 2019.

More than one participant commented that they felt an increased familiarity with the sonification over the course of the day and that this made it easier to understand the different districts. They also highlighted that in real-world situations, operators would undergo specific training for the sonification, something they deemed extremely useful. We consider both comments very valuable because, as the reader might recall, we intentionally chose not to give the participants extended training to allow a freer exploration of the prototypes and foster the development of personal strategies in the relationship with the sonification. Last but not least, one participant stated that the information contained in the sonification was more than enough if combined with visualizations that the operator could access for further analysis.

Structure of the sonification

In general, the structure of the sonification was well received. As Figure 61 illustrates, participants found that both the minimum (one second) and the maximum duration of the sonification (10 seconds) were appropriate for conveying meaningful information without intruding on the work routine. Three repetitions were also deemed appropriate ("The first attracts my attention, the second allows me to understand, with the third I confirm my understanding"), with the suggestion of cutting it down to only one repetition in case of regular behavior.

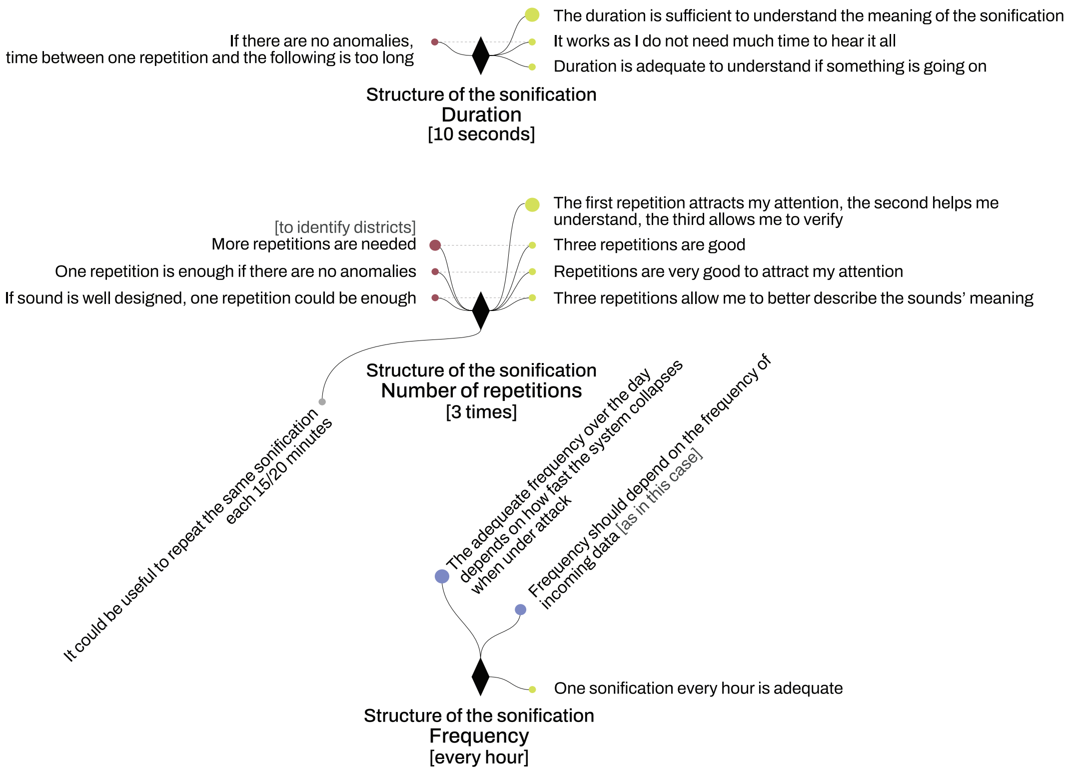


Figure 61. Results of qualitative research: structure of the sonification. Adapted from Terenghi 2019.

In terms of the frequency of the sonification (one ever hour) the testers, experts in digital/physical systems and cyber-security, confirmed what the team at iTrust had already stated: that the frequency of the alert depends on the reaction of the system under study during an attack i.e., on how fast the system could collapse due to the consequence of an intrusion. In the case of water networks, one hour seem to be an appropriate time frame.

Application in a real-world context

When asked about the usage of the sonification in the real world, all participants answered positively. As Figure 62 recaps, the reasons stated in support of the adoption of sonification as a real tool for monitoring anomalies in the context of water infrastructure included: it requires a low cognitive effort; alerts reach you fast; you can use it even when you are busy with other tasks and it leverages an alternative sensory modality to vision. Issues that need to be solved for real world adoption include the fact that the sonification can be missed if you are on the phone, leave your desk or are interacting with colleagues. The need to be close to the source of the sound (which partly explains why the user can miss the sonification) is another reason mentioned by the participants.

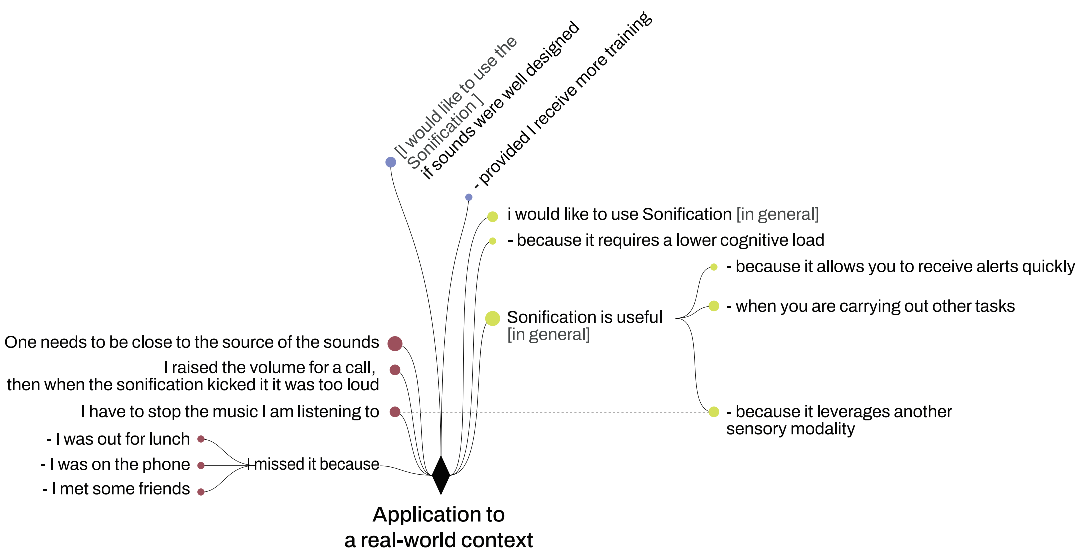


Figure 62. Results of qualitative research: application in a real-world context. Adapted from Terenghi 2019.

The need for the sonification to co-exist with other sources of sound, such as music operators might listen to through the same headphones or the occasional web call, was also raised. More in general, with this being extremely relevant to this thesis, some testers highlighted that for a real-world usage they would expect sounds to be "well designed": in particular, sounds should not be too repetitive; they should be familiar enough to remain at the periphery while still attracting attention when needed; they should not intrude on the normal work routine. Some of the participants added that they would request more training before considering adopting the sonification in a real-world scenario.

Comparative analysis of the four scenarios

In Figure 63, feedback on each of the four scenarios is clustered by gathering comments that emerged during different moments of the interviews, notes that the participants added to the spreadsheet during the quantitative experiment and answers to the direct question: “Which scenario did you find more intuitively understandable?” and “Would you be able to pick a scenario in which you liked the sound most?”. Results are displayed in two parts: the upper part of Fig.63 shows comments on the data-to-sound mapping and the type of sounds. The lower part gives more detail on individual preferences and focuses on feelings of like/dislike, intuitiveness and emotional response to the sound.

As can be seen from the analysis, none of the scenarios emerged as a clear ‘winner’, with feedback being inconsistent throughout the qualitative material we analyzed. For the mapping strategy ‘duration’ (scenarios 1), comments ranged from positive (“It’s easier to recognize the anomaly level”) to negative (“I cannot distinguish the districts because sounds superimpose”). Scenarios 2, based on repetition, seem to collect more preferences (“It is more interesting and pleasant”; “The distinction among districts is easier to grasp because sounds repeat over time”, “It is easier to recognize the anomaly level”). On an opposite note, other participants stated that the ‘repetition’ strategy prevents an easy recognition of the anomaly level as it lacks a comparative scale against which to measure the frequency of the repetition. With regard to tuned (version A) and non-tuned (version B) sounds, a wider use base is certainly needed to achieve significant results. As we can see in figure 63, comments are extremely subjective and range from “I prefer non-tuned sounds” to “I prefer musical sounds” or “Non tuned sounds annoy me”. In general, scenario 2, in both versions, seem to obtain more positive results. Talking about Scenario 2A participants use words such as “the best”, “the most pleasant”, “my favorite” or even “it makes me happy”. Scenario 2B is described as “easier to understand because sounds are different”, “It sounds like a game”, “It’s fun”, “I used to attribute a personal meaning to each sound”. Scenarios 1 seemed to be less attractive, in general, with scenario 1A being referred to as “OK”, while scenario 1B (duration with non-tuned sounds), seemed

How to read

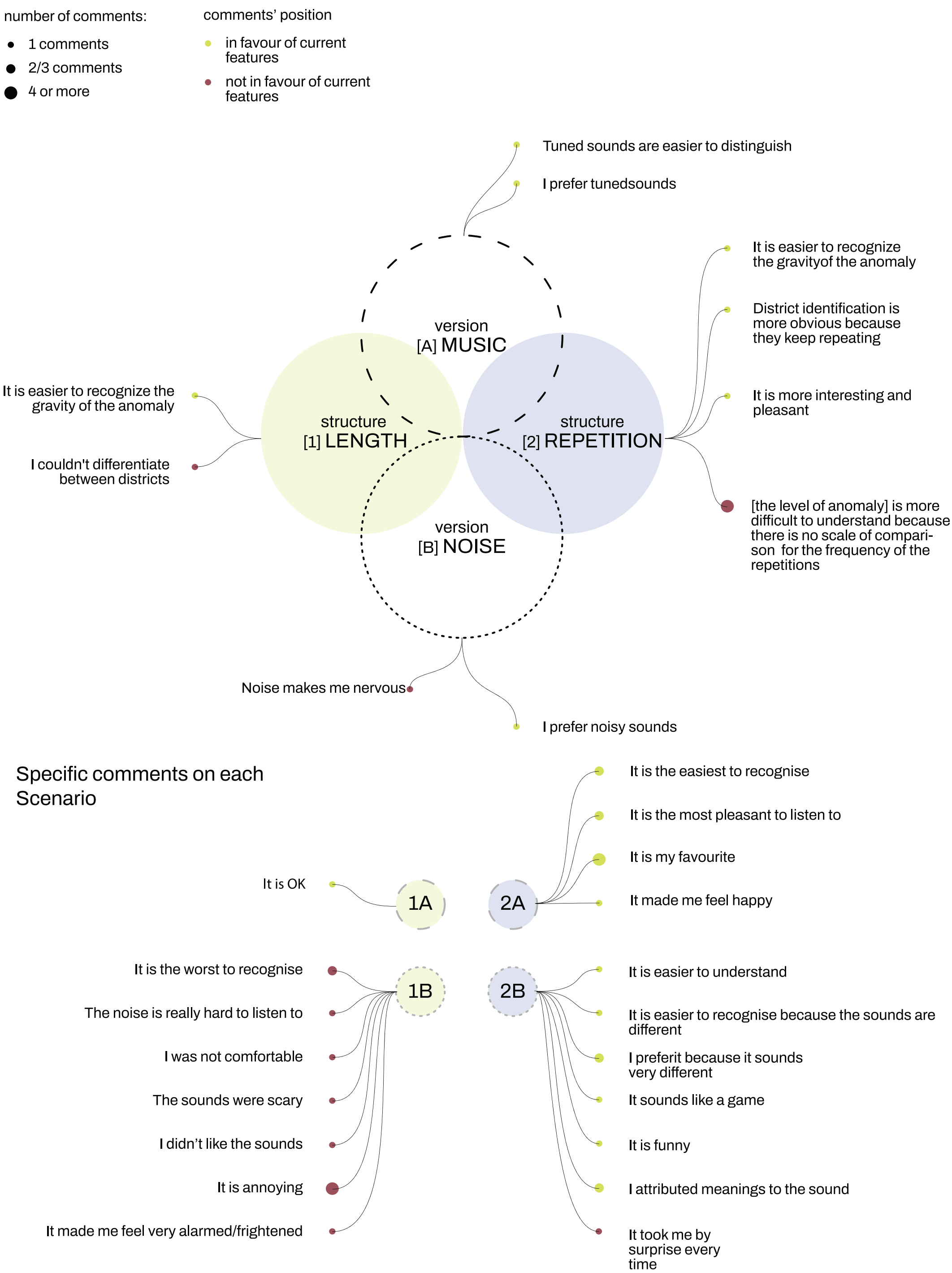


Figure 63. Results of qualitative research: comparative analysis of the four scenarios. Adapted from Terenghi 2019.

to collect all the negative feedback. Participants refer to it as “the worst”, “very difficult to listen to”, “It makes me feel uneasy” and even “It sounds terrible”, “It’s annoying”, “It frightens me and upsets me”. On an integrated perspective, it is interesting to compare results from the quantitative test with such comments. Participants seemed to assume that they performed better in the scenarios they favored and worse in those they disliked. When presented with their individual results, they were quite surprised to realize that that was not always the case: in particular (see Fig.51) scenario 2B scored very poorly in the identification of the anomaly level of the network despite participants judging it the easiest to understand, whereas the “terrible” scenario 1B performed better even in the most difficult task of identifying the anomaly level for each district (see Fig.55).

User journey- We asked the testers to describe the actions they undertook during testing days in relation to the sonification. As mentioned, we did not provide them with training sessions but only, through the website, with the individual sound keys to each district and the mapping strategy for each scenario. Some of the participants retrospectively highlighted how they would have benefited from training with feedback. In particular, they imagined a training based on exercises similar to those taken during the experiment but with an immediate feedback on errors. From the interview it is possible to gauge how participants made up for the lack of training through some steps which are meant to help them in evaluating their understanding of the sonification, such as listening to the key-sounds both before and after the sonification. It also emerged that they felt more comfortable with the sonification as time went by and as they developed their personal strategies to decode the sounds and get insights on the behavior of the network.

In Chapter 6, we will compare the experimental results of the first DA with those of the second, with the goal of highlighting similarities which could inform the design of other sonifications for the monitoring of anomalies in digital and digital/physical networks.

CHAPTER 5 —
Secondo tema, seconda variazione.
Real-time anomaly detection in Internet networks.

Soundtrack. *T. Zé. The return of Tom Zé.*

The second design action was developed during a 12-month internship I carried out at Instituto 'Ibermática de Innovación – i3B', the R&D unit of the Spanish company Ibermática[28]. At the time I joined, a small team of researchers, experts in Artificial Intelligence, was involved in a publicly funded project dedicated to the exploration of Machine Learning – supported cyber-security applied across a number of different scenarios, among which Internet services, financial services, healthcare and Industry 4.0. The main goal of the project, which ran from 2017 to 2019 under the acronym SUCESO[29] was the "creation of a collaborative and shared database of Indicators of Compromise (IOC)[30] for each of the involved sectors, with the goal of alleviating the effects of the main threats (known and not yet known) to Internet networks. In particular, ML techniques were applied to identifying patterns of anomalous behavior that was unknown to date as well as to facilitate the interpretation of this by non-expert users through visual analytics tools." (Hazitek 2018, translated by Sara Lenzi).

28.
https://ibermatica.com/

29.
Machine learning SUpported CiberSEguridad. The acronym plays with the Spanish word SUCESO which means 'occurrence', 'incident'.

30.
"Indicator of compromise (IoC) in computer forensics is an artifact observed on a network or in an operating system that, with high confidence, indicates a computer intrusion" (Wikipedia)..

In the landscape of the Spanish industrial ecosystem, the project aimed to introduce a series of innovative approaches. In particular:

- While IOC are mainly used in computer forensics investigations to identify the causes of a cyber-attack after the attack has taken place, the project focused on the development of a real-time solution for the monitoring of cyber-attacks.
- The solution would be flexible and scalable enough to be applied across the partners' sectors: Industry 4.0, IT and finance, healthcare.
- The project would target domain experts of the sectors involved rather than data analysts or cyber-security experts.

Figure 64 offers an overview of the project's architecture. The infrastructure of SUCESO was built using data collected from the different partners. A main anomaly detection algorithm was designed and trained on this data. After that, the system was applied to specific use cases defined by the project's partners.

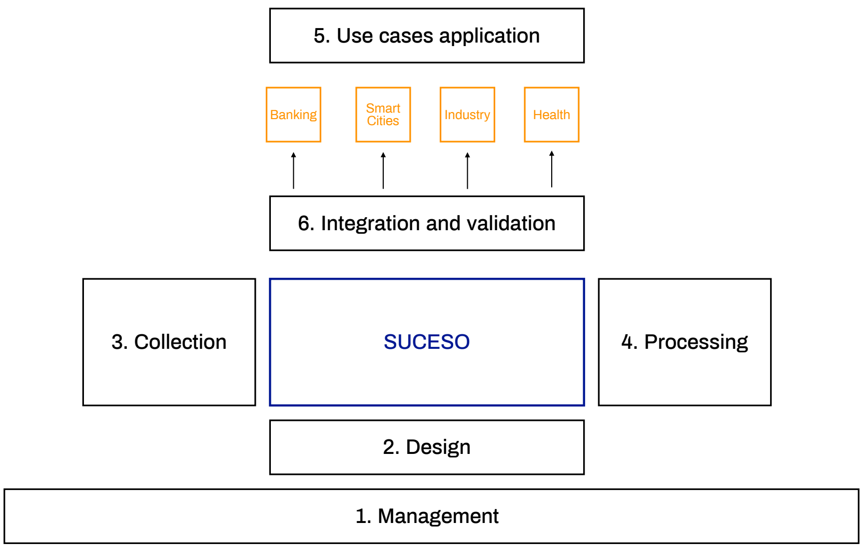


Figure 64. Architecture of SUCESO. Memoria técnica de Proyecto 2018.

When I joined i3B in October 2018, the project was reaching a milestone for which Ibermática was responsible. In this phase, the data mining infrastructure and the anomaly detection algorithm were designed and implemented. The team developed PLAGEMODA (PLAtform for the manaGEment and MOdeling of DAta), a tailor-made platform for data collection, storage and analysis with an attached dashboard for visual analytics. The anomaly detection algorithm was designed using deep learning techniques and was trained with a database of about 7000 historical data from the Internet network of Ibermática.

A schema of the solution as of October 2018 is shown in Figure 65: the platform collects data from a digital network. Data are managed, analyzed and presented on a dashboard to be consulted by the operators of a SOC for routine tasks. At the same time, data are streamed to the algorithm which has been purposely designed to identify anomalies due to cyber-crime. Data are then extracted from the algorithm's log file and visualized for forensic analysis.

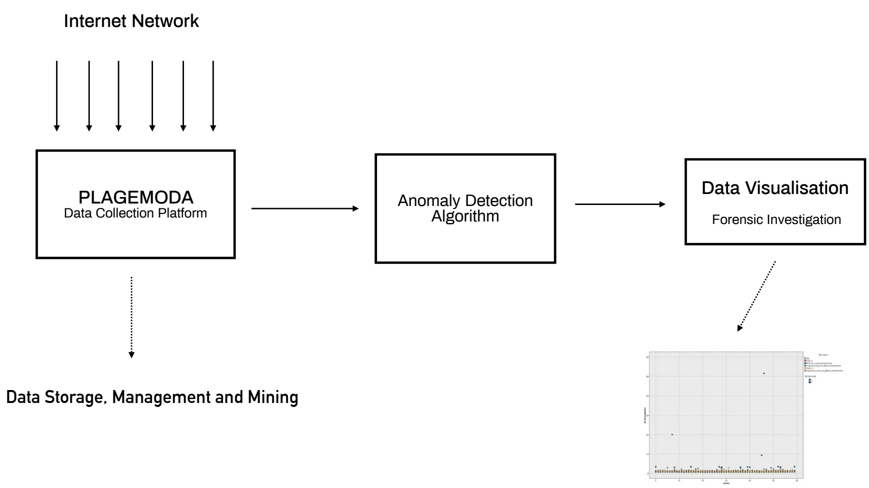


Figure 65. Scheme of SUCESO's solution.

One of the main focuses of the project was the relationship "between the real-time data analysis and the data visualization tools to allow for an optimization of decision-taking through the usage of efficient and intuitive visual analytics." (Memoria de Proyecto cit., p.63, trad. Sara Lenzi). Moreover, these visualization tools would be flexible enough to be easily understood by a diversity of users i.e., domain experts from sectors as different as banking and healthcare. In late 2018 the development of the visualization tools was focused, as mentioned above, on presenting data for – mainly – forensic analysis. In this sense, the visualizations were not (yet) meant to be used in real-time, but rather as a tool for the analysis of an attack after its occurrence. There was no mention of sonification in the “Memoria Técnica de Proyecto” (cit.), that nonetheless dedicated more than twenty pages to a description of the state of the art in data representation techniques for the purpose of analysis, which also includes tactile interfaces. To my surprise, there were no specific plans to integrate a dedicated alarm system to alert the operator of an incoming threat.

The team tables the idea to use sonification to represent the data filtered and interpreted by the anomaly detection algorithm with the goal of informing the operator of the behavior of the network at all times and alert in the event of anomalies. This tool and the process of conceptualizing, designing and implementing it is the object of the second Design Action. Figure 66 shows how the sonification tool was integrated in the existing architecture of SUCESO. Similarly, to the first DA, sonification was explicitly meant to support, not replace, the visualization systems.

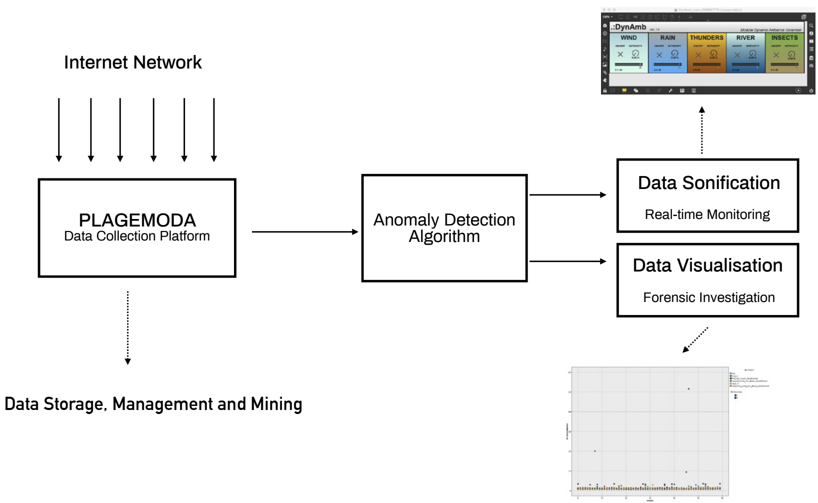


Figure 66. The SUCESO solution with the integration of data sonification.

This chapter will describe the different work phases of the DA. After defining the use case, we explored the dataset, the choice of sound types and the mapping strategies through the creation of a demo that was shared with the team during an internal event. Once the overall concept and the choice of sound types were validated, we moved forward with a detailed analysis of the dataset and the mapping strategy, while working on integrating the sonification app with the existing anomaly detection system. The first prototype of the sonification tool was released in this phase. In the third phase, I was presented with the opportunity to test the prototype in a real environment, the simulation laboratory of Tecnalia[31], a large public-funded research institution in Bilbao, Spain and partner of Ibermática for cyber-security projects. A second iteration of the prototype was developed based on the Tecnalia use case. An experimental protocol was put in place to validate the new prototype within Tecnalia's facilities and with the company's personnel. As I will describe in the following paragraphs, plans for the experimental phase had to be downgraded due to the continuing restrictions due to the COVID-19 pandemic. Nonetheless, I will be able to share with the reader the results of both quantitative and qualitative experiments and draft plans for the future development of the sonification tool, which is still ongoing.

31. https://www.tecnalia.com/en/

SUCESO: Machine learning SUpported CyberSEcurity

Ibermática is a Spanish company headquartered in San Sebastián, Basque Country. It is one of the country's largest IT provider for businesses and, as we might expect, it is currently entering the market of AI-based solutions for cybersecurity. Its main clients are large institutions such as regional and local authorities, private and public healthcare providers, national associations (such as La Once, the Spanish national association for the visually impaired), utilities (such as Internet and electricity providers) and so on. Within the company, i3B is the unit dedicated to the exploration of future services and products. The unit works mainly within the framework of publicly funded projects (by the European Union or by national and regional funds). A group of about twenty professionals and researchers, i3B is also home to a smaller team specifically working on potential applications of AI. The team's mission is to explore and expand the range of AI applications and is not preoccupied with immediate commercial take-up. Therefore, a certain margin of free experimentation is allowed and even encouraged - exploring the use of sonification as a tool for the real-time monitoring of anomalies in the prevention of cyber-attacks on digital networks certainly fell within this mindset.

Within the scope of SUCESO, the AI unit of i3B was tasked with the following responsibilities, illustrated in Fig. 65:

- To design and develop the data collection, storage and mining platform PLAGEMODA.
- To design and implement the anomaly detection algorithm that would be installed on PLAGEMODA and would interpret the incoming data through its models.
- To design the visualization tools, later developed in visualization (for forensic analysis) and sonification (for real-time monitoring).

Anomaly detection and hierarchical modeling - In Chapter 4, I described how anomaly detection algorithms in digital-physical systems work: they mainly compare the current behavior of the system to an ideal regular (i.e., with no anomalies) behavior on which the AI model is trained. When the current behavior diverges from the expected ideal behavior, the algorithm would flag this situation as 'anomalous'. In the digital networks of SUCESO, the model is trained to identify Indicators of Compromise (IoC) i.e., traces that harmful code, used to penetrate and take control of digital systems, leaves behind. The creation of a shared, trans-sectorial database of known IoCs as well as unknown IoCs (which would be predicted by the AI model) was actually the main goal of the project. The preliminary training of the model was to be carried out on a purposely made digital platform within the Ibermática network, where a database composed of historical logs of about 2M data was loaded. Based on a preliminary analysis of recurring, typical IoCs found on the database, four use cases that the model would be trained to identify were defined:

1. Detection of foreign Ips.
2. Numerous contacts with the same service within a fixed time-frame.
3. Same IP attacks on the same service with different methods (POST, GET, DELETE, etc.).
4. Same IP attacks on the same service with different clients (Chrome, Mozilla, Safari, etc.).

Given the large amount of data the model would have to manage to concurrently identify these four cases, the SUCESO team opted for a layered approach in which the algorithm was trained to prioritize the identification of some cases over others. The information received by the algorithm was structured hierarchically, with the identification of IP assigned to layer 'Zero' and information related to the behavior of the different services of the network to sit on the more specific layer 'One'. For each layer, a different anomaly detection model was designed. Consequently, as shown in Figure 67, the use case 'detection of foreign IPs' would sit on Level 0 while the remaining three cases would sit on Level 1. In an iterative process, knowledge on recurring patterns of attacks extracted by the model at Level 0 will inform knowledge on Level 1. Conversely, improved knowledge on anomalous behavior of the services in Level 1 would be generalized to refine the model sitting on Level 0, so that the AI system continued to learn while minimizing calculation and data management.

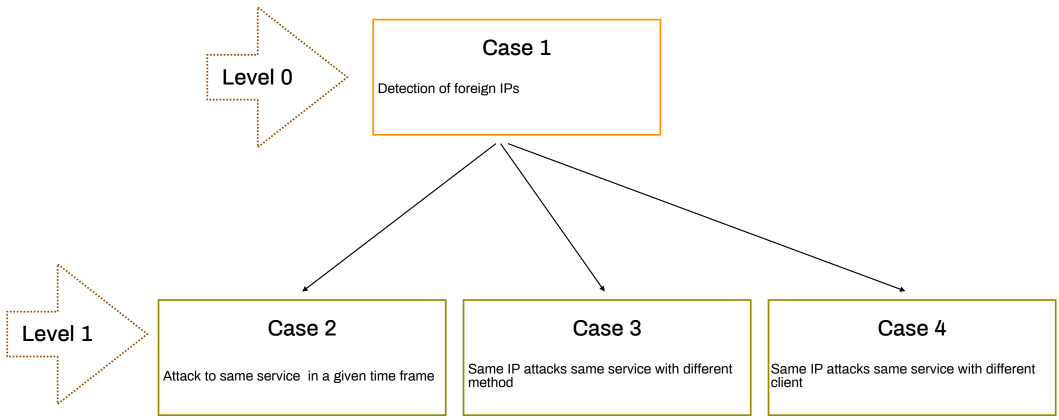


Figure 67. Hierarchical organization of the anomaly detection algorithm. Memoria técnica de Proyecto 2018.

We decided to focus the sonification design on Level 0 for various reasons. First of all, this would be the first layer to be implemented in the anomaly detection model. Secondly, and similarly to the first DA, we conceptualized the sonification as an early detection tool that would get alerts on global anomalies in the overall behavior of the network rather than receive detailed information. The operator would subsequently refer to other sources of analytical information (visualization, log files, etc.) for further analysis.

Characteristics of the data set- Table 3 shows how the raw dataset of 2,048,442 entries of historical data from the Ibermática network appear, before being filtered by the anomaly detection model. The relevant fields used for modeling the anomaly detection algorithm based on the four cases described above are:

Orig_h: source IP

Orig_p: source port

Resp_h: destination IP

Resp_p: destination port

Method: access method

Host: destination host

Uri: uniform resource identifier

Referrer: Uri reference

User_agent: access client

Request_body_len: length of the request

Response_body_len: length of the response

Campo	Medida	Valores	No se encuentra	Comprobar	Rol
orig_h	Nominal	"192.168.202.100","192.168.202.101",...		Ninguno	Entrada
resp_h	Nominal	"192.168.201.2","192.168.202.112","1...		Ninguno	Entrada
Año	Continuo	[2012,2012]		Ninguno	Entrada
Mes	Continuo	[3,3]		Ninguno	Entrada
dia	Continuo	[16,17]		Ninguno	Entrada
dia_semana	Continuo	[6,7]		Ninguno	Entrada
hora	Continuo	[12,22]		Ninguno	Entrada
minuto	Continuo	[0,59]		Ninguno	Entrada
mehod_Count_transformed	Continuo	[-0.12080067572814189,42.39281493...		Ninguno	Entrada
uri_Count_transformed	Continuo	[-0.12080067572814189,42.39281493...		Ninguno	Entrada
user_agent_Count_transformed	Continuo	[-0.12080067572814189,42.39281493...		Ninguno	Entrada
request_body_len_Mean_transformed	Continuo	[-0.07552786285612918,54.53271992...		Ninguno	Entrada
response_body_len_Mean_transformed	Continuo	[-0.022712103556862904,77.7829436...		Ninguno	Entrada
Record_Count_transformed	Continuo	[-0.1208006757281 42,42.392814931 3...		Ninguno	Entrada

Table 3. SUCESO raw dataset.

orig_h	resp_h	Year	Month	Day	Day/week	hour	minute	method	uri	user_agent	request_body_len	response_body_len	Record
192.168.202.79	192.168.25.203	2012	3	16	6	19	45	16,238	16,238	16,237981	-0,075528	-0,022146	16,23798
192.168.202.79	192.168.25.203	2012	3	16	6	19	46	42,3928	42,3928	42,392815	-0,075528	-0,022146	42,39282
192.168.202.79	192.168.25.203	2012	3	16	6	19	47	23,2783	23,2783	23,278311	-0,075497	-0,021891	23,27831
192.168.202.79	192.168.25.103	2012	3	17	7	19	28	-0,1196	-0,1196	-0,119584	18,732591	0,022433	-0,11958
192.168.202.96	192.168.24.152	2012	3	16	6	16	11	-0,1099	-0,1099	-0,109854	2,468665	-0,020779	-0,10985
192.168.203.63	1,92168E+11	2012	3	16	6	17	50	8,39368	8,39368	8,39368	-0,075528	-0,022712	8,39368
192.168.203.63	1,92168E+11	2012	3	16	6	17	51	8,18002	8,18002	8,180017	-0,075528	-0,022712	8,180017
192.168.203.63	1,92168E+11	2012	3	16	6	17	52	8,49828	8,49828	8,498281	-0,075528	-0,022712	8,498281
192.168.203.63	1,92168E+11	2012	3	16	6	17	53	7,97446	7,97446	7,974463	-0,075528	-0,022712	7,974463
192.168.203.63	1,92168E+11	2012	3	16	6	17	54	8,29354	8,29354	8,293538	-0,075528	-0,022712	8,293538
192.168.203.63	1,92168E+11	2012	3	16	6	17	55	8,33773	8,33773	8,33773	-0,075528	-0,022712	8,33773
192.168.203.63	1,92168E+11	2012	3	16	6	17	56	8,25786	8,25786	8,25786	-0,075528	-0,022712	8,25786
192.168.203.63	1,92168E+11	2012	3	16	6	17	58	9,22928	9,22928	9,229276	-0,075528	-0,022712	9,229276
192.168.203.63	1,92168E+11	2012	3	16	6	17	59	8,41679	8,41679	8,41679	-0,075528	-0,022712	8,41679
192.168.203.63	1,92168E+11	2012	3	16	6	18	0	8,20759	8,20759	8,207587	-0,075528	-0,022712	8,207587
192.168.203.63	1,92168E+11	2012	3	16	6	18	1	8,60978	8,60978	8,609775	-0,075528	-0,022712	8,609775
192.168.203.63	1,92168E+11	2012	3	16	6	18	2	8,18894	8,18894	8,188937	-0,075528	-0,022712	8,188937
192.168.203.63	1,92168E+11	2012	3	16	6	18	3	8,43503	8,43503	8,435034	-0,075528	-0,022712	8,435034
192.168.203.63	1,92168E+11	2012	3	16	6	18	4	7,8877	7,8877	7,887701	-0,075528	-0,022712	7,887701
192.168.203.63	1,92168E+11	2012	3	16	6	18	6	7,89459	7,89459	7,894593	-0,075528	-0,022712	7,894593
192.168.203.63	1,92168E+11	2012	3	16	6	18	7	8,23921	8,23921	8,23921	-0,075528	-0,022712	8,23921
192.168.203.63	1,92168E+11	2012	3	16	6	18	8	7,66958	7,66958	7,669578	-0,075528	-0,022712	7,669578

Table 4. Raw data at Level 0.

Table 4 shows the same data aggregated at the hierarchical layer

'Level 0'. The columns show each of the relevant fields obtained from the network's raw data before being interpreted by the anomaly detection algorithm.

At Level 0 the goal of the algorithm is to identify both incoming and outgoing foreign IPs. The so-called 'granularity' of the database (the minimum unit of information at which the analysis is performed) is therefore fixed at the semantic couple IP of origin/IP of destination, as shown in the first two rows of Table 4. The frequency at which the granularity is considered is one minute. All the other parameters of the network are aggregated based on the granularity and the frequency.

Applying the anomaly detection model- As described in Chapter 4, anomaly detection algorithms do not focus on storing information on the irregular or uncommon behavior of a system. On the contrary, they store information on the regular behavior of a given system in order to detect any deviations from the norm – even unknown deviations. In fact, in the specific context of cybersecurity, threats are extremely diverse as new attacks are being carried out with new means all the times, making it particularly challenging to keep an updated list of existing threats. It would not be efficient or sustainable for an anomaly detection algorithm to try to include all new threats in its database. Indeed, a good anomaly detection system must focus on refining knowledge on a network's regular behavior. This allows an algorithm to detect – and with time, prevent – attacks, including ones carried out with unknown means. Figure 68 shows how the anomaly detection algorithm used in SUCESO works. Within the network, elements are clustered within 'peer groups' based on the similarity of their behavior. Elements of the network are assigned to the peer group with which they share most characteristics. The algorithm then measures the distance from the center of the peer group (which represents the ideal normal behavior) for each incoming data point. The distance (represented by the arrow in Fig.68) informs us of the deviation from the norm of each element of the network under study at any given time.

Within this overarching structure, at Level 0 the anomaly detection algorithm for SUCESO considers only the index of anomaly for the whole network every minute (Table 5, column I) and the reasons that produce such anomalous behavior, clustered around the so-called 'fields'. The algorithm identifies three fields (Table 5, columns K, M, P) which contain details of the specific reason causing the anomalies (for example in Field 1, column K, the reason is the . Fields are hierarchically ordered: 'field 1' represents the main reason for the anomaly, while fields 2 and 3 identify less serious reasons that may contribute to, but not determine, the global anomaly index. The 'field impact' (Table 5, columns L, N and P) indicates the relative weight of each field measured in the overall anomaly index.

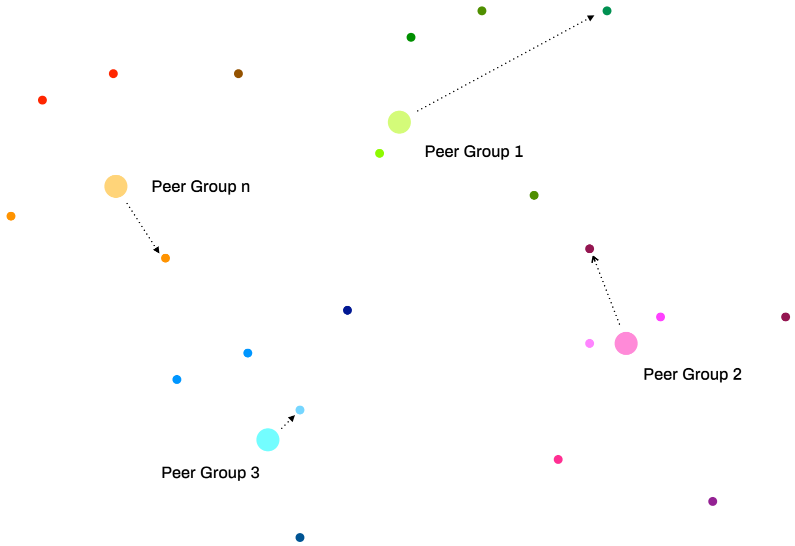


Figure 68. The network's elements interpreted by the algorithm.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
orig_h	resp_h	Year	Month	Day	Hour	Minute	$O-Anomal	$O-AnomalyInde	$O-PeerGrou	$O-Field-1	$O-FieldImpact-	$O-Field-2	$O-FieldImpact-	$O-Field-3	$O-FieldImpact-3
192.168.202.102	192.168.23.101	2012	3	16	17	49	F	0,822747	2	resp_h	0,547936	orig_h	0,218324	minuto	0,149336
192.168.202.103	192.168.25.202	2012	3	16	17	49	F	0,800384	2	resp_h	0,409121	orig_h	0,350722	minuto	0,153509
192.168.202.102	192.168.21.103	2012	3	16	17	49	F	0,79413	2	resp_h	0,531731	orig_h	0,226192	minuto	0,154717
192.168.202.110	192.168.229.156	2012	3	16	17	49	F	0,692256	2	orig_h	0,38876	resp_h	0,33371	minuto	0,177486
192.168.202.76	192.168.229.156	2012	3	16	17	49	F	0,687265	2	orig_h	0,384453	resp_h	0,336133	minuto	0,178775
192.168.203.63	192.168.229.101	2012	3	16	17	50	T	3,472138	1	Record_Count	0,168438	method_Count	0,168438	uri_Count	0,168438
192.168.202.102	192.168.23.202	2012	3	16	17	50	F	1,594561	2	resp_h	0,268564	Record_Count	0,123954	method_Count	0,123954
192.168.202.97	192.168.28.252	2012	3	16	17	50	F	1,531803	2	orig_h	0,534593	resp_h	0,335908	minuto	0,084232
192.168.202.79	192.168.26.203	2012	3	16	17	50	F	1,215592	2	resp_h	0,580521	orig_h	0,256208	minuto	0,106143
192.168.202.79	192.168.24.252	2012	3	16	17	50	F	1,066523	2	resp_h	0,521889	orig_h	0,292019	minuto	0,120979
192.168.202.79	192.168.26.152	2012	3	16	17	50	F	1,028938	2	resp_h	0,504426	orig_h	0,302686	minuto	0,125398
192.168.202.79	192.168.26.253	2012	3	16	17	50	F	1,015935	2	resp_h	0,498151	orig_h	0,30656	minuto	0,127003
192.168.202.79	192.168.26.202	2012	3	16	17	50	F	0,99727	2	resp_h	0,488695	orig_h	0,312297	minuto	0,12938

Table 5. Layer 0 database.

Additionally, as can be seen in column H (Table 5), the 'anomaly index' in column I is coupled with the indication of anomaly 'true/false'. In fact, unlike the algorithm of DA1, the SUCESO algorithm is programmed to make a judgement based on its knowledge of the regular behavior of the network. In other words, it discriminates between an anomaly level that signals a real attack and an anomaly level due to non-malicious deviations from an ideal ‘perfect’ behavior that is rarely found in real-world situations.

The visualization tools- As previously mentioned, visualizations were designed to support forensic investigations on a cybersecurity incident. The example in Figure 69 shows one minute of data where the operator can identify the level of anomaly for each of the different causes in field 1 (the length of the request, the length of the response, the day of the week in which the activity happens, the source of destination IP) as well as whether the anomaly is considered ‘true’ i.e., due to cyber-attacks, or ‘false’ by the system.

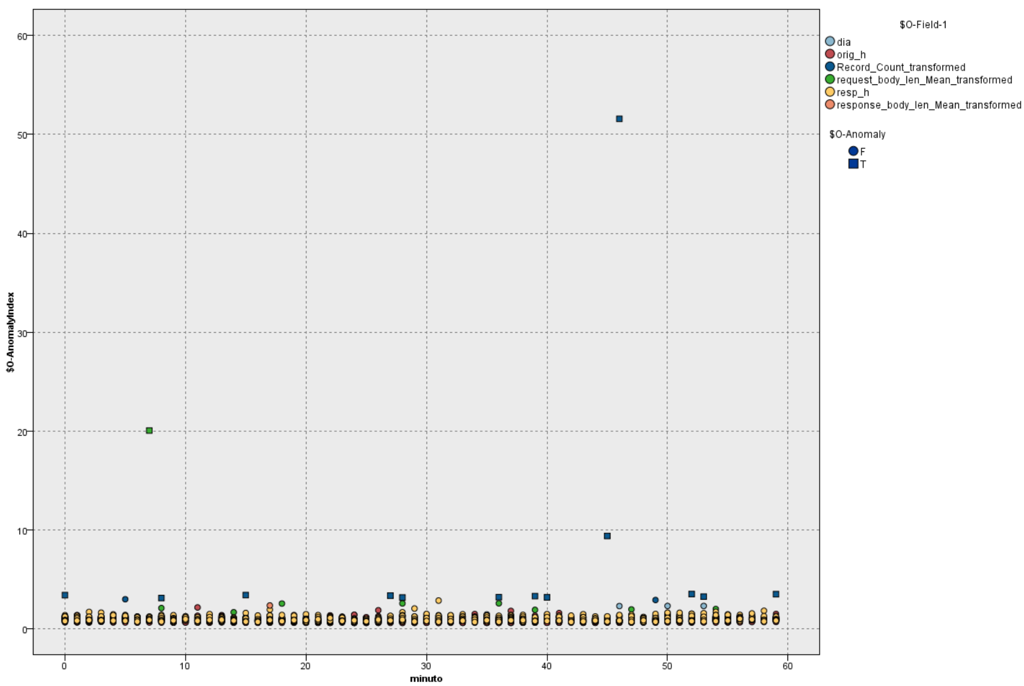


Figure 69. Excerpt of visualization for Field 1.

Together with the i3B team, we decided to focus the design of the sonification on complementing the information provided by the visualization. The sonification would alert the operator in real-time on the global anomaly index of the network, on the importance of each field and on the status of the anomaly (true or false). The visualization would instead allow the operator to retrieve analytical information on the specific cause of the attack and its exact value as well as the value of each of the elements of the network extracted from data shown in Table 4.

Data sonification design - Research

The design process for the sonification included two phases. In the first phase, I investigated possible mapping strategies and types of sound to push the boundaries of the information contained in the dataset, with the goal of producing a 'demo' sonification to share with the team in order to gather some quick feedback. In the second phase, the design process was structured to develop a fully functioning prototype for testing in a real environment.

Definition of the use case- The use case for this project was defined in line with existing literature and findings from the first DA. On the one hand, and as described in Chapter 4, the large amount of visual information which the operator of a SOC is exposed to during a typical workday and the consequent risk of overload of the visual channel pave the way for the introduction of an alternative sensory modality to convey additional information, such as alerts on cyber threats. Additionally, sound is believed to possess several characteristics that make it a suitable representation means in this specific context: it sits at the periphery of our attention until needed, allowing for the user to carry out other, mainly visual, tasks; it occupies a different sensory channel, thereby avoiding additional information overload; it leverages the human capability to easily recognize recurring patterns or sudden changes in continuous sound events. On the other hand, results obtained in the experimental phase of the first DA had clearly shown that domain experts considered favorably the possibility of using auditory representations of data in a real work environment. The same users highlighted the need for sound to be able to integrate with the work environment as well as with the tastes, needs and habits of the users. From this previous research and from brainstorming sessions with the Ibermática team, I refined the definition of the use case and added two main constraints for the design of the sonification:

- Following the experimental results of the first DA, we decided that the sonification would be optimized to be played through headphones. Indeed, I had learnt that operators tend to work with headphones, either for listening to their favorite music or for sheltering from environmental noise.
- Unlike digital-physical networks, where data are collected from physical sensors within a time window that heavily depends on

the system under study (water plant, energy grid, and so on), in fully digital networks, such as the Internet networks of the second DA, data are collected in real-time. For this reason, the sonification had to be designed as a continuous soundscape created by data streaming in real-time.

Other interesting points had emerged during the first DA follow-up interviews. In particular:

- Users would welcome a background sound that is non-intrusive and potentially relaxing, as long as its design characteristics do not interfere with the analyst's capacity to focus on their tasks.

- Additionally, users seem to have a tendency to attribute a familiar meaning to concrete sounds in order to facilitate identification and recollection. This seemed to be coherent with a listener's tendency to attribute a semantic or a causal meaning to sounds (Chion 2015, cit.) i.e., to consciously or subconsciously interpret sounds they hear in relation to a code or to the source of the sound.

- There seems to be a certain expectation on the aesthetic aspects of the sound experience, with users seemingly unable to differentiate between aesthetic experience and the cognitive appreciation of a particular sonification: in the first DA, some of the experts thought it was easier to understand the sonifications they aesthetically preferred. Participants had also pointed at how aesthetic considerations and feelings of pleasure encouraged them to look for insights from the data.

We decided to conceptualize the sonification as a data-driven natural soundscape, in particular the soundscape of a forest. Listening to a forest soundscape, the operator would make sense of the behavior of the network and its components by leveraging a simple metaphor: the Internet network as a forest in which the individual components (for instance birds, insects, and the sound of leaves) contribute in shaping the forest's overall behavior.

We hypothesized that this soundscape would help the user:

- To locate herself in a familiar sonic environment, thus minimizing the need for musical or audio training to distinguish and interpret different sounds.

- To limit the cognitive effort required to monitor the sonification while dealing with other tasks. Indeed, the peripheral monitoring of natural soundscapes is a common acquired experience through which humans are able to effortlessly navigate everyday life (Lakoff & Johnson cit.).

Additionally, literature seemed to support the use of natural soundscapes to increase the efficiency and effectiveness of process-monitoring tasks (Hildebrandt, Hermann, Rinderle-Ma 2016; Hildebrandt, Hermann, Rinderle-Ma 2014, cit.; Debashi and Vickers 2018; Vickers et al cit.). In particular, Hildebrandt and colleagues applied a data-driven soundscape approach to real-time process monitoring in industrial and manufacturing contexts. Their study concludes that natural sounds "as encountered everyday a thousand times would be attractive and yield better compatibility with long uses", while other type of sounds "would not meet the acceptability threshold for extended (e.g. full day) use" (Hildebrandt, Hermann, Rinderle-Ma 2015, cit., p.195). On a similar note, Vickers and colleagues state that "we are already used to dealing with everyday background sound and quickly deciding what sounds need attending to and what sounds can be pushed to the attentional background. A soundscape offers the sonification designer the potential to leverage this innate information processing capacity in such a way that important changes in the cyber environment become salient in the soundscape." (cit., p.5).

Demo: exploring mapping and metaphors- Similarly to the approach used for the first DA, I decided to produce a quick demo to present the team with a 'feeling' of the natural soundscape and at the same time freely test some data-to-sound mapping options, before digging thoroughly into the database. From the main dataset of about 7000 entries, I selected a few lines (corresponding to 15 minutes of data) which included both regular and anomalous data. A set of natural sounds, which included recorded samples of birds, insects, other animals like frogs, rain, thunder and wind, was sourced from my personal sound library.

In the previous section I described the main features of the database at the hierarchical level we were working at (see Table 4). To recap, the sonification was meant to represent the 'anomaly index', 'anomaly true/false' and the three 'fields' with their individual 'field index'. In this first demo, in the spirit of exploring the database in search of alternative points of view, I attempted a different clustering of the data.

I asked the Ibermática experts to cluster the parameters associated with the couple source IP/destination /IP - the minimum unit of information at Level 0 - around higher-level semantic descriptors. As shown in Table 4, the algorithm identifies, for each line of information, the reasons for anomalous behavior within each field. For instance, the first line of column K tells us that the primary reason for the anomaly is the source IP (with an associated field impact of 0,547936). For the purpose of the demo, the dataset was re-arranged in order to associate each parameter in the fields to one of the semantic descriptors defined by the experts. Figure 70 shows a list of these descriptors and the elements they contain.

SEMANTIC GROUPS	
Operation Content	Request_body_len (length of the request) Response_body_len (length of the response)
Access Identification	Orig_h (source IP) Orig_p (source port) puerto origen Resp_h (destination IP) Resp_p (destination port)
Access Method	Method of access
Operation Destination	Host (destination host destino) Uri (destination access service) Referrer (URI reference)
Access Client	User_agent (access client)
Operation Status	Status_code (request status)

Figure 70. Semantic descriptors.

In the demo I used a different sound to represent each semantic descriptor. Figure 71 shows a matrix of the mapping which also serves as a score for the demo of the sonification. Six different natural sounds were employed, one for each semantic cluster. The weight of each anomaly (the field impact, in the original dataset) was used to control the volume for each sound. Similarly to the first DA demo, data on the field impact was normalized on a 1 to 5 scale, where 1 represented the lowest impact and 5 the highest impact of the specific cause of the anomaly, on the whole network. The network's overall anomaly index was represented by the sounds of an incoming storm (also normalized on a 1 to 5 scale from the lowest to the highest gravity with the anomaly), with the thunder used as an indicator of anomaly 'true'.

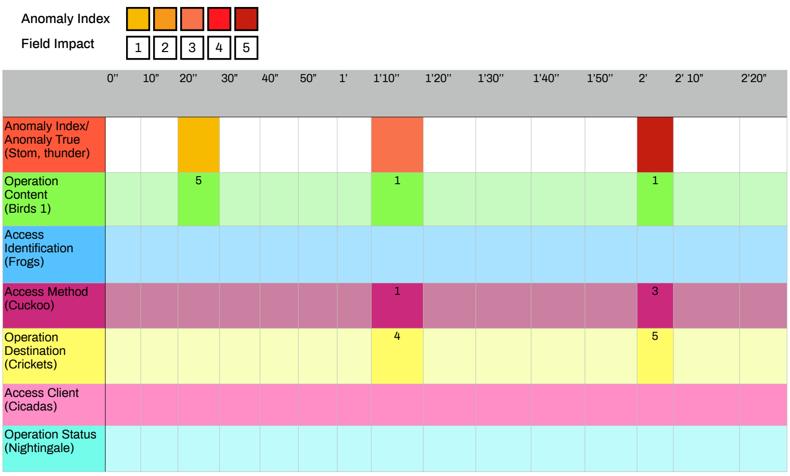


Figure 71. Mapping matrix and demo score.

As shown in Fig.71, a short sound sequence of about two minutes was obtained from an excerpt of 15 minutes of data showing both anomalous and non – anomalous behavior, in order to explore the auditory result of both scenarios.

The sequence can be listened to at the following link:
https://www.saralenzi.com/design-actions-material PSW: SonificationDesign2021

The demo was presented at an internal session of i3B where different research teams shared ongoing projects for peer feedback. Feedback was extremely positive on the usage of a natural soundscape but less optimistic on the usage of six different sounds to account for the semantic clustering, plus a seventh sound for the representation of the anomaly index and even an eighth sound for the anomaly 'true'. In the team's opinion, with eight different sounds to remember and distinguish, the amount of cognitive effort risked overriding the efficiency of using a natural soundscape i.e., a sound material which did not require specialized training. As we will see in the following section, I therefore decided, for the prototyping phase, to revert back to the original categorization in three fields, anomaly index and anomaly 'true', thus reducing by almost half the need for different sounds. The use of volume as an indicator of the anomaly level was also debated, as it was argued volume carries a subjective, non-controllable component influenced by the listening environment, the quality of the device used and a personal sensitivity to loudness. As we will shortly see, for the design of the prototype, we radically changed the rules of how adjustments in the dataset determine changes in the soundscape, moving away from volume towards concepts such as the density of the sound material. Last but not least, the scaling of the level of anomaly and the impact of the fields along a pre-determined scale of gravity also gave rise to an interesting discussion. While on the one hand, this choice allowed me to design a simple correspondence between data and sound that was easy to understand for the user and easy to implement from a design perspective, on the other, it introduced a potential bias in the interpretation of the algorithm's results. In fact, the anomaly detection algorithm employed in SUCESO only distinguishes between a true and a false anomaly. It is the direct responsibility of the operator - a responsibility the sonification is meant to facilitate, not to override - to interpret, by leveraging contextual information on the behavior of the whole network, the gravity of a 'false' anomaly or even, perhaps, identify a false positive (which, as we saw in Chapter 4, is still a rather common occurrence in AI based anomaly detection). Therefore, in the design of the prototype the pre-determined scale was removed in favor of a linear representation of the anomaly index coming from the algorithm, thus leaving the final decision on the gravity of the anomaly to the human operator.

The sonification prototype- The second phase of the sonification design started with a recap of the use case i.e., the description of the user and the context of usage, as well as the definition of the goals of the prototype:

- The sonification was to complement existing visualization tools. In particular, the sonification would be used as a real-time, continuous monitoring system located at the periphery of the user's attention, ready to move to the center if needed, whereas the visualization would be an 'on demand' analysis tool.
- The user would presumably listen to the sonification through headphones during working shifts of several hours.
- From a sound design point of view, the prototype would be

based on a natural soundscape, which had received a very positive feedback during the demo session.

- From the perspective of the data to sound interaction, experts had suggested moving away from the use of volume to indicate changes in the behavior of data: a new type of mapping strategy had to be explored.
- From a technical point of view, the prototype would have to interact with the data platform, in order to receive a data stream from the network and interpret this through the anomaly detection algorithm.

Time was dedicated to reviewing the dataset in light of feedback from the demo session. To recap, the anomaly detection algorithm receives data from the PLAGEMODA platform (see Figure 65) with a one-minute frequency. At the hierarchical layer we are working on, Level 0, data granularity is defined by the couple 'source IP/destination IP'. This means that all data collected from the network are grouped on the basis of a unique combination of IPs and clustered over a one-minute timeframe (see Table 4 for more details). This specific use of time, with clustering of real-time within a one-minute frame, combined with the specific granularity at which we are working, is problematic when working with sonification. In fact, as all data arrive at the same moment (for instance, in the first four rows of Table 4, four different IP couples are stamped with the same time, 17:49), we have virtually no way of assigning a different 'arrival time' to each string of data. At the same time, we want to avoid associating a different sound to each row i.e., to each combination of IPs. This would not only exponentially increase the quantity of different sounds the user has to recognize but, more importantly, it would be of no use in terms of the main goal of the sonification: to help the user detect and react to anomalies. For the first prototype, we decided to temporarily overlook the issue of the specific approach to time in the dataset and focus on anomaly detection.

Anomalies are represented by the algorithm through fields. Fields tell us, at every minute and for each IP couple, which element of the network is responsible for the anomalous behavior. The network presents, in fact, an index of anomaly at all times. These anomalies are not always of a malignant origin i.e., due to cyber-threats or cyber-attacks.

As the chief of the project within i3B put it “Networks are as anomalous as life itself is. There is never a situation in life that fully adheres to an ideal model of perfection, glitches are everywhere”. To measure how much anomalous behavior impacts the overall status of the network, fields are weighted i.e., not only are they in hierarchical order, with field 1 representing the main cause of anomaly and field 2 and 3 representing secondary causes, but, also, each field is assigned an index, which is indicated in the dataset as field impact. Based on the behavior of the fields, a global anomaly index is assigned by the algorithm to the overall network. Only if this level is understood by the algorithm as malicious (i.e., it moves beyond a certain threshold, established by the algorithm itself), an indication of anomaly ‘true’ is also issued to the operator.

Unlike the context of the first DA, we received a continuous stream of data (every minute): we therefore needed a continuous soundscape instead of punctual sound events played every hour as in the first DA. Based on this consideration, existing literature and users’ feedback from both the previous DA and the demo’s expert session, we drafted a design concept where the digital network is represented by an evolving natural soundscape, in particular, the soundscape of a forest. We understand the forest as a metaphor for the Internet network: forests, like digital networks, are dynamic systems, in which the behavior of individual, independent actors has consequences on the behavior of the overall system. Through this metaphor, the user/listener makes sense of the behavior of the network and its components, by interpreting the behavior of the forest both as a whole and in its individual elements (birds, insects, wind through the trees, crackling of wood, animals moving through the bushes, rain and the occasional clap of thunder). We expect the listener to be able to switch from an analytical listening experience, in which she would focus on details of the behavior of an element of the forest to gauge details on the behavior of the corresponding element of the network (for instance, she would listen to how birds are behaving to get insights on how field 1 is behaving), to a global listening experience, in which she focuses on what the whole forest is telling her: a storm coming, with wind increasing and rain getting stronger, is an indicator of something serious happening to the network.

After defining the embracing metaphor, I moved on to sketch the interaction paradigm in detail. Table 6 gives an overview of the data-to-sound mapping for each data category and introduces what I called the 'chaos factor', the main rule that determines the behavior of sound over time.

Type of Data	Description	Type of Sound	Sound behavior
Field 1	Main reason for the anomalous behavior.	Birds	Sound is always on; its behavior over time is controlled by the field impact.
Field 2	Secondary reason for the anomalous behavior.	Insects	Sound is always on; its behavior over time is controlled by the field impact.
Field 3	Least important reason for the anomalous behavior.	Noises of the forest (wind through the leave, steps through the bushes, wood crackling…)	Sound is always on; its behavior over time is controlled by the field impact.
Field Impact 1,2,3	Index of the relative weight of each field on the global anomaly of the system.	It controls the 'chaos factor' each field.	'chaos factor' is directly proportional to the value the field impact.
Anomaly Index	Global index of anomalous behavior in the system.	Rain	Sound is always on; its behavior over time is controlled by the 'chaos factor'.
Anomaly Yes/No	Detection of anomaly due to cyber-attack.	Thunder	Alarm triggered when anomaly = true. When anomaly = false there is no sound.

Table 6. Data to sound mapping matrix

As we can see from Table 6, two mapping strategies coexist in the sonification: one controls, through the so-called chaos factor, the behavior of the different sound categories (birds, insects, noise and rain) over time. The other triggers an individual sound (thunder) in a very specific circumstance: when the algorithm decides the anomaly is malignant i.e., it is due to a cyber-attack.

After defining the mapping strategy, we moved on to take a series of decisions related to the specific design of the sound. Figure 72 recaps the choices made for the first iteration of the prototype. The schema includes choices on which data are represented by which sounds under the tags 'sound' and 'sound design' and which specific rules govern how data and sound evolve over time, under the tag 'effect'. As mentioned above, two different mapping strategies are applied to the four categories of the field impacts and the anomaly index and the indication of anomaly 'true'.

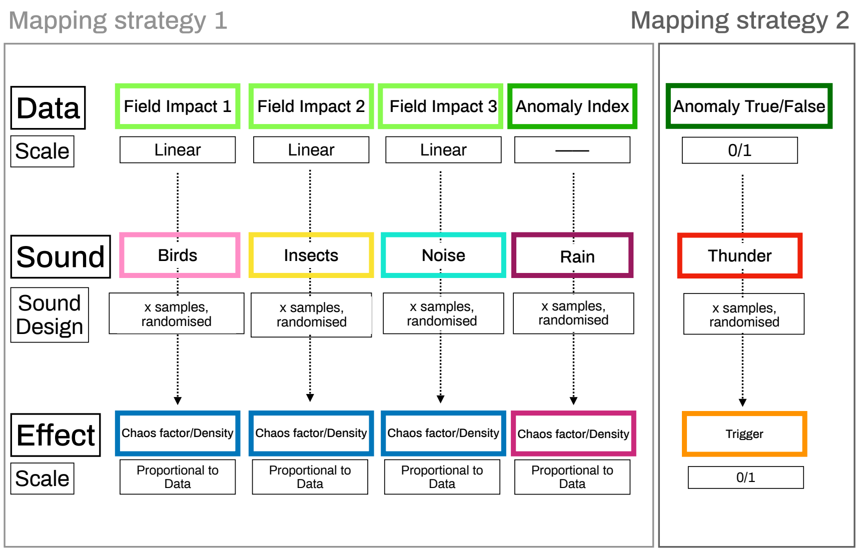


Figure 72. First prototype data-to-sound mapping matrix.

In the following sections I will describe in detail the two mapping strategies and how data and sound relate, in particular through the chaos factor.

Mapping strategy 1: communicating the network behavior- The first mapping strategy aims to communicate to the operator the global behavior of the network, at all times and unobtrusively i.e., as a peripheral monitoring support.

Sounds of birds, insects, forest noises and rain are mapped to, respectively, field 1, field 2, field 3 and the anomaly index. The reader will recall that the relative importance of anomalies in field 1 is higher than that in field 2 and 3. I hypothesized that birds tend to attract the attention of human listeners more than insects and generic noise. Therefore, birds were used to represent the most important field 1, while insects and lastly, noise, were associated with the progressively less relevant fields 2 and 3. Soundscape studies do not, to the best of my knowledge, attribute a hierarchy of importance to different elements of a soundscape. Research on the role of emotions (and notably, on the impact of familiar and known sounds) in the perceptual evaluation of the soundscape are ongoing (Fiebig, Jordan, Moshona). Studies have been run on the role of this emotional recognition of a soundscape in order to distill guidelines for the design of sounds (Asutay et al. 2012). The recent lockdowns of human populations in most of the countries of the world have brought to the attention of the media the perceived increase in bird sounds (Lenzi, Sádaba, Lindborg 2021). Researchers have tried to evaluate the growth in natural sounds in urban contexts demonstrating that no relevant increase in the presence of birds in our cities can be proven (Gordo et al. 2020). Research has also focused on the iconic and cultural importance given to birdsong by human listeners (Lenzi, Sádaba, Lindborg cit.). Considerations on the iconic role of bird calls supported the decision to use birds as the most prominent sound category in the sonification of DA1. On the other hand, conversations with Ibermática experts confirmed that valuable insights might come if the operator were to gain a global perspective on how the three fields interact, rather than focusing only on the field of primary importance. While maintaining the hierarchical structure of the fields in the organization of the sound categories, although, I wanted to leave the door open to subjective interpretation in the listener, in the hope that this would, in the longer term, add precious insight to the interpretation of the network behavior.

I assumed that the network keeps oscillating along a continuum from order (when anomalous behavior in the network is very low) to chaos (when anomalous behavior in the network increases). In the majority of cases, chaos represents the norm i.e., it is not due to malicious activity, but rather normal deviations from an ideal model. As suggested during the demo's expert session, I decided that the sonification would not inform the operator of the gravity of the anomaly i.e., we did not embed a pre-determined level of anomaly (for instance, on a scale 1 to 5). In fact, the operator is supposed to gather this information from the acoustic behavior of the system. It is one of the assumptions of the mapping strategy that, over time, the operators will develop a listening proficiency that will enable them to distinguish even subtle changes in the behavior of the soundscape, just as we are all able to distinguish subtle changes in our everyday experience of real-world soundscapes.

The core of the first mapping strategy is what I called the 'chaos factor'. Leveraging the analogy between distortion (i.e., an anomaly in the sound signal) in sound as an anomaly in data, we defined the chaos factor as an auditory element meant to alter the perception of the listener and alert her on anomalies in the network. In other words, the chaos factor is what informs the operator that "something is not right". The chaos factor is proportionally mapped to data i.e., when data show a higher anomalous behavior, the behavior of the corresponding sounds becomes more chaotic.

Figure 73. Sketching exercise for the 'chaos factor'.

In the first phase of the sonification design, I investigated several options, drawing inspiration from both the literature and sound design practices. Figure 73 shows how I used a visual sketch to brainstorm on a possible association between emotional dimensions, visual imagery and acoustic representations. This sort of preliminary cross-modal exercise is common practice in my compositional activity. In this case, the exercise turned into a family 'workshop' during which we listened to music, talked about the meaning of the selected emotions and improvised graphic representations. Even if it might sound far-fetched to directly connect the exercise with the later selected sound effects for the data-to-sound mapping, I think it is worth sharing with the reader as an example of the multi-faceted ways through which tacit knowledge makes itself explicit. In the sketch, emotional dimensions analogous to the opposition between calm and chaos were listed, including regularity vs irregularity, positive vs negative, safety vs danger, peace vs stress, familiar vs unfamiliar, normality vs abnormality, certainty vs uncertainty.

Then I proceeded to hypothesize how these opposite extremes could be represented through the evolving sounds of a forest. Options included: a progressive distortion of the sounds; a progressive deconstruction of the sound through (for instance) delay, bandpass filters, noise gate or granular filters; an increase in the density of sound samples belonging to the same category; an increase in the intensity of the sound samples belonging to the same category.

In the current version of the prototype, we designed the chaos factor as a proportional increase in the density of sounds. We define 'density' as the progressive accumulation of sound samples within the same category. This choice was also supported by the results of the first DA, in which the mapping strategy based on repetition (scenario 2), where sound samples progressively accumulated within the same time frame at the increase as the anomaly, seemed to give better results and was positively valued by the testers. As I will describe shortly, in the prototype, each sound category has an associated folder in which numerous, short samples are stored. The number of samples played within a specific time frame i.e., how 'dense' that particular sound category is, is a decision taken by the sound processing engine of the prototype based on the incoming value of the anomaly for each category.

Mapping Strategy 2. Alerting on an incoming cyber-threat- The second mapping strategy uses sonification as a prompt alarm system rather than a global monitoring system . As described earlier, the approach of a storm in the sonification, with the rain sound getting more intense, signifies an increase in the level of the anomaly. A clap of thunder is triggered by the sonification tool when anomalous behavior is recognized as malicious by the algorithm. The second strategy, therefore, reports to the operator the decision made by the algorithm on whether the system is being attacked. As such, it functions on a binary, 0-1 dichotomy: the attack is either happening or is not happening. In the event of an incoming message of 'anomaly true' the sonification will trigger the sound of thunder, thus attracting the attention of the operator who will have to enforce a procedure which involves a series of pre-determined steps (accessing the visualization to analyze specific data, run checks on specific areas of the network, escalate the alarm to colleagues and so on). Unlike the first mapping strategy, the goal of the second was to remain at the center of the operator's attention. The operator must not miss information on an incoming attack. Note that the alarm alerts the operator when the attack has already occurred – it triggers a reaction, but it does not help prevent an attack. This would be the role of the first mapping strategy.

Conclusions- A digital tool was designed to implement the mapping strategies in real-time. The tool receives data from the anomaly detection algorithm and sonifies these according to the design decisions described in this Chapter. This prototypal tool was later used to test the initial hypothesis and specifically, that the sonification:

- Met the requirements of blending with the context (an office working environment), was pleasant to hear and intuitively understood, minimized cognitive effort to decode information, leveraged familiar sounds with a high sound quality.
- Allowed for peripheral monitoring.
- It had the potentialities, in the long run, to help operators predict anomalous behavior, thus reducing errors due to false positive in the anomaly detection algorithm.

Data sonification design – Prototype

We implemented the prototype on a local network using Python[32] and Max/MSP[33] by Cycling '74. A Python script simulates data streaming from the PLAGEMODA platform to the prototype, while a Max/MSP patch translates it into sounds. The two tools communicate via the OSC (open sound control) protocol[34]. For each data category, with the exception of the anomaly 'true', which I will describe shortly, the sound engine accessed an individual folder (one per category) where a number of samples (currently about twenty per folder, each about two seconds-long) are pre-loaded. Figure 74 shows a simple diagram of the prototype's architecture:

32. https://www.python.org/

33. https://cycling74.com/products/max/

34. http://opensoundcontrol.org/osc

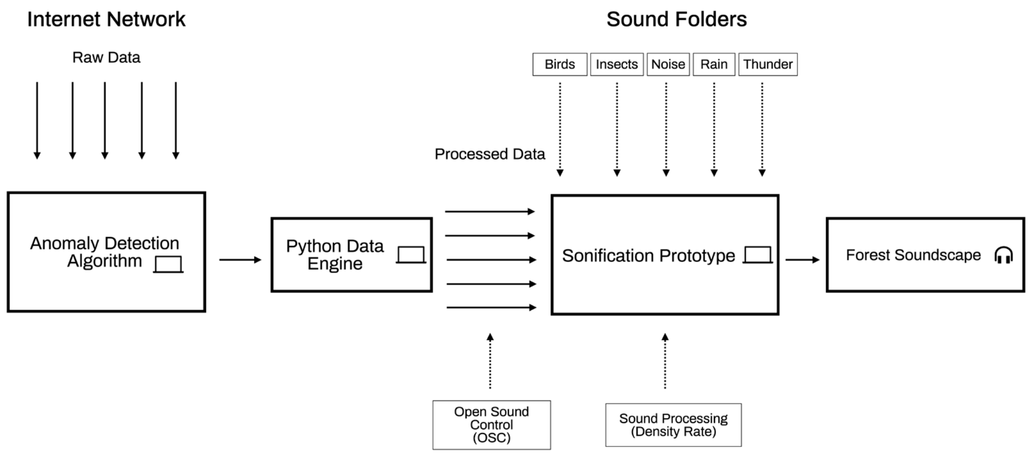


Figure 74. Diagram of the prototype's architecture.

Most of the sounds were sourced from my personal sound library, an archive of several hours of soundscape recordings taken over the years from locations mainly in Italy and South East Asia. For this version of the prototype, samples from the tropical rainforests of Malaysia, Indonesia and Thailand were selected. A smaller set of samples was sourced from the archive of ‘Sounds of the Forest’ by Timber Festival UK[35], released under the Creative Common license ‘share- alike’. All samples and the respective authors are cited in the Appendix.

35. https://timberfestival.org.uk/soundsoftheforest-soundmap

The reader might recall that we had temporarily set aside an issue related to the specific management of time by the anomaly detection algorithm. At the granularity Level 0, the algorithm clusters data from the network at one-minute intervals. Therefore, the database has multiple entries (of different IP couples) within the same minute. There is no way we can know in which order, within the minute, data on a specific IP couple was received. After consulting experts at Ibermática and experts in network security on various hypotheses, we decided to establish an artificial order: data would be scaled within each minute, by dividing the minute by the number of incoming data for the same field. For instance, if the database showed four strings of incoming data from ‘field 1’ at minute 10.25:00, data from ‘field 1’ would be sonified every 15 seconds from 10.25 to 10.26. In the event of an attack, data regarding the anomaly would have priority, so that the prototype would play the sound corresponding to the anomalous data before the others and for the entire minute. This solution was judged to be a good compromise since, despite forcing the dataset into an artificial order, it would not interfere with the overall goal of the sonification i.e., to detect anomalies in time to activate an appropriate reaction to avoid damage to the network.

The interface - Figure 75 shows the prototype interface in its second iteration. The five data-to-sound categories are displayed in the main building blocks of the interface. Each category has an on/off control through which the operator can select and isolate a specific sound group for closer listening. A volume control allows for real-time adjustments while providing a visual cue on the behavior of the sound (at a higher density the volume of the sound category will presumably be higher). A knob shows in real-time the percentage of the density rate on a scale 1 to 100, for instance, in Fig. 75 Field 3 has a density rate of

15.2%. This also provides an additional visual cue on the behavior of data: at a higher anomaly level, the knob shows a higher percentage of density. The four grey blocks under the categories ‘anomaly index’, ‘field’ 1, 2 and 3 display additional information. In the first line, the exact value of the incoming data is displayed (in Fig.75, field 3 has a value of 29.780001).

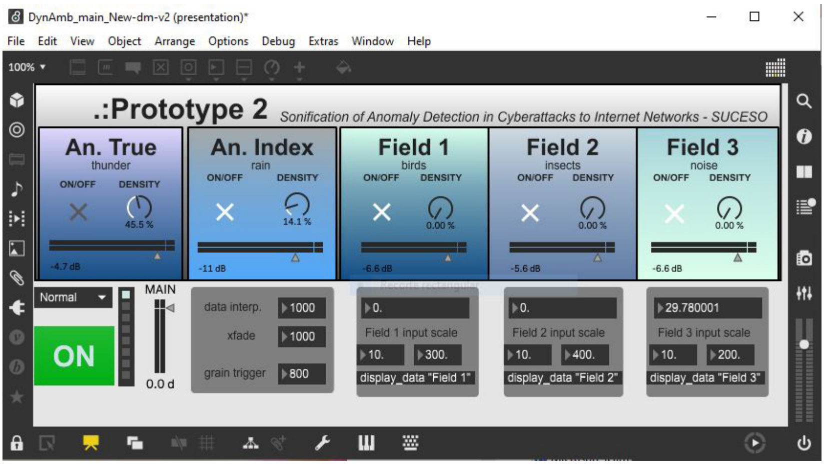


Figure 75. DA2 sonification prototype, v.2.

The input scale for each field i.e., the range in which sound samples are accumulated (which conversely determines the speed at which the 100% density is reached) can also be specified. This way, the listener can adjust the perceived density of samples if she feels the accumulation is too evident or not sufficiently evident. Other controllable parameters (all of which are expressed in milliseconds) include: the data interpolation value i.e., the time interval at which new incoming data blend with the data currently sonified; the crossfade range i.e., the time interval at which two different samples merge; the grain trigger i.e., the timeframe at which the application receives a new trigger to refresh its operations. The grain trigger is of particular importance as it determines the density percentage: at a higher trigger value, less sounds will play within the given timeframe. The ‘anomaly true’ block is activated with the sound of thunder when an anomaly is identified as true by the algorithm. A number of samples of thunderclaps of different quality, intensity and spatial arrangement are pre-loaded into the corresponding folder. At the moment, samples are triggered based on an accumulation principle: when receiving a ‘true’ value from the data simulation engine (see Fig.76), the prototype triggers the first sounds of the corresponding folder. If the following incoming value is normal the sound will play until the end and no further sounds will be triggered. If the prototype receives another 'true' value, the following sample (for example, thunder_sample02) is triggered while the first sample is still playing. The engine keeps playing the samples in order until 'thunder_sample05' and then starts again from the first sample of the list until the anomaly value returns to 'false', thus creating a polyphony of thunderclaps. To avoid repetitions that could be recognized by the user, samples are triggered randomly, and a pitch distortion ratio is also applied to mask the repetition. A dropdown menu under the 'anomaly true' block (displaying 'normal' in Fig.75) turns to 'attack' when the anomaly is 'true', thus providing an additional confirmation, should the operator misinterpret the sound in such a key situation. The red button 'ON' is the main power button for the sound of the prototype and can be switched off at any moment by the operator.

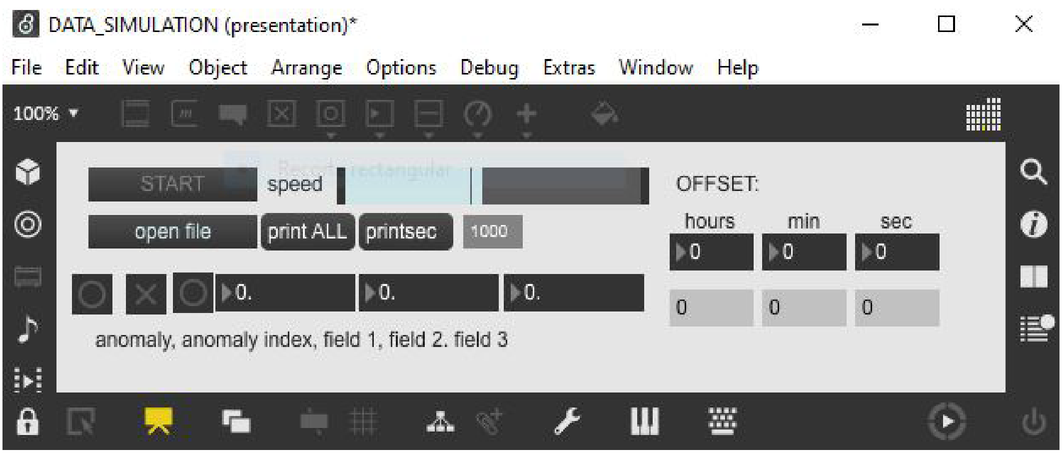


Figure 76. Data simulation engine.

The 'data simulation' display (Fig.76) is used when the prototype is not connected to a real-time streaming data engine. It was programmed for fast testing with different datasets. As the following section describes, the data simulator was extensively used as the main source of real-time data during the experimental phase, when restrictions due to the COVID-19 pandemic severely limited our access to physical laboratories and therefore, the possibility to develop and test the prototype with a real data streaming platform. The data simulator displays the values for each incoming data as they are received, in the following order: ‘anomaly true/false’, ‘anomaly index’, and the value for the ‘field impact’ of field 1, field 2 and field 3. In Fig. 76, the data string tells us that the values for each field are: field 1 23064.98, field 2 30.8 and field 3 26.18. When users open the interface they have the option of selecting an external database (as a .csv file) through the button ‘open file’. The ‘speed’ cursor allows the user to decrease or increase the reading speed of the file (in milliseconds) i.e., to extend or shorten the one-minute timeframe of the database.

To the right of the interface, the user has the option to set the starting time, which refers to the time of the database. This option was integrated for the experimental phase as we imagined the participant might need to stop the application and start again from the same point at a later stage (for example, after a short break). The ‘print’ commands export a log of all the data received. This button was conceived for ease of consultation on historical data. At any point in time the operator (or the participant in the experiment) might want to have a closer look at data that has already been sonified.

Test Bed: Tecnalia

In the last phase of the SUCESO project, the possibility emerged to test the prototype in a real-world scenario in the premises of the Spanish research institution Tecnalia, in collaboration with the team in charge of the investigation on cross-industry anomaly detection systems for cyber-security. The team extracted about a week of data from their local network, a rather small network composed of seven IP addresses. Raw data was then interpreted by the SUCESO anomaly detection algorithm at the same hierarchical layer used to build the prototype, Level 0. Finally, we arranged a session of experiments at the premises of Tecnalia, where a lab simulating a Security Operations Center is used to test new applications in a near real-world scenario.

The dataset- Some adjustments had to be made to the prototype in order to process data from the Tecnalia network. Both the raw data and the dataset interpreted by the algorithm presented a few differences from Ibermática’s corresponding data. Table 7 shows a sample of the raw data obtained from the Tecnalia local network. As we can see from Table 7, the network only collects a small amount of information: at the granularity of sourceIP/destinationIP, we were only informed on the type of protocol used to communicate and the length of the information exchanged. Table 8 shows a sample of the Tecnalia database interpreted by the anomaly detection algorithm, with both regular ('normal') and anomalous values ('outliers').

DATE	avg_packet_length	ip_dst	ip_src	protocol
2020-09-09T17:00:00.000Z	641.44	192.168.2.9	192.168.2.12	mms
2020-09-09T17:00:00.000Z	274.45	192.168.2.9	192.168.2.12	tcp
2020-09-09T17:00:00.000Z	673.45	192.168.2.10	192.168.2.12	mms
2020-09-09T17:00:00.000Z	673.45	192.168.2.10	192.168.2.12	tcp
2020-09-09T17:00:00.000Z	237.49	192.168.2.9	192.168.2.11	tcp
2020-09-09T17:00:00.000Z	672.62	192.168.2.9	192.168.2.11	mms
2020-09-09T17:00:00.000Z	96.0	192.168.2.9	192.168.2.11	ntp
2020-09-09T17:00:00.000Z	402.39	192.168.2.10	192.168.2.11	mms
2020-09-09T17:00:00.000Z	216.88	192.168.2.10	192.168.2.11	tcp

Table 7. Raw data from the Tecnalia local network.

DATE	ip_src	ip_dst	browser	cotp	data	mms	ntp	tcp	_tpkt	label	outliers	distance
2020-09-18T10:25:00	192.168.2.10	192.168.2.12	0	0.0	0	138.0	0	74.15	0	0	Normal	20.52
2020-09-18T10:25:00	192.168.2.10	255.255.255.255	0	0.0	0	0.0	96	0.0	0	2	Normal	39.85
2020-09-18T10:25:00	192.168.2.11	192.168.2.10	0	0.0	0	325.67	0	236.7	0	0	Normal	21.82
2020-09-18T10:25:00	192.168.2.11	192.168.2.9	0	0.0	0	532.14	96	254.87	0	2	Normal	26.18
2020-09-18T10:25:00	192.168.2.11	192.168.2.93	0	82.0	0	118.89	0	69.0	0	0	Outlier	23064.98
2020-09-18T10:25:00	192.168.2.93	192.168.2.11	0	76.0	0	152.78	0	60.38	0	1	Outlier	19829.17
2020-09-18T10:26:00	192.168.2.93	192.168.2.11	0	65.0	0	153.0	0	60.0	0	1	Outlier	14513.53
2020-09-18T10:28:00	192.168.2.11	192.168.2.93	0	82.0	0	223.0	0	69.0	0	0	Outlier	23063.69
2020-09-18T10:28:00	192.168.2.93	192.168.2.11	0	76.0	0	245.0	0	61.5	0	1	Outlier	19828.24

Table 8. Dataset interpreted by the anomaly detection algorithm.

Some changes in the dataset parameters can be identified. Indeed, when applied to a different network with different raw data, the algorithm returns slightly different parameters within the same conceptual schema. Lines from 'browser' to '_tpkt' show the value of each communication protocol (the only parameter collected from this specific network, as shown in Table 7). 'Label' indicates the group to which each IP couple with its specific communication protocol in that specific minute, pertains. Groups (or 'labels') are determined by the algorithm itself and represent clusters of similar behavior. For every group, the algorithm defines an ideal normal behavior represented by the center of the group (previously illustrated in Figure 68) from which, from time to time, elements of the groups might deviate. The amount of this deviation (in Table 8, 'distance') tells us the amount of anomalous behavior for that particular combination of IPs during a specific minute.

To remind the reader, through the value of ‘fields’, the previous database informed us of the three main reasons for the global amount of anomalous behavior in the network. Due to the limited size of the Tecnalia network, experts at Ibermática set the algorithm to provide information, instead, on the distance from normality of each element of a group at a given time. In order to minimize changes to the prototype and test its scalability when adapted to a new network, we created a matrix of correspondences (see Table 9) where we adapted existing functionalities to the new database. Note that the number of groups is automatically determined by the algorithm but can also be decided manually by the data scientist. In this case, we ran several tests with different amounts of groups: twelve – as automatically determined by the algorithm - five, three and two. Results confirmed that, whatever the number of groups, the same anomalies were found in the same groups ‘label 1’ and ‘label 0’. This supported the decision to cluster data around three labels in order to use the same three corresponding sound categories of the sonification prototype (birds, insects and noise).

	SUCESO	-> -> -> TECNALIA	SOUND DESIGN
1	Field 1	Label 1	Birds
2	Field 2	Label 2	Insects
3	Field 3	Label 3	Forest’s Noise
4	Anomaly True/False	Outlier/Normal	Thunder
5	Anomaly Index	----	Rain

Table 9. Correspondences between the SUCESO and the Tecnalia databases.

Table 9 summarizes the correspondence between the two datasets. ‘Labels’ takes the place of ‘fields’, with ‘field impact’ replaced by the value of the distance from the center of the group. ‘Labels’ are sonified by changes in the density parameter in the corresponding sound categories ‘birds’, ‘insects’ and ‘noise’. The indication of ‘outlier’ and ‘normal’ replaces ‘true’ and ‘false’ and is sonified by thunder. As we can see, in this case the database does not give a value for the global ‘anomaly index’, due to the extremely reduced size of the network. Still, we decided to maintain a soft sound of rain as a continuous forest background so that the operator would still hear ‘something’ in the absence of incoming data. This would help avoid a crucial misunderstanding:

that the absence of sound could be confused with the absence of data rather than a malfunctioning of the sound system. Figure 77 shows the original mapping strategy updated for the testbed:

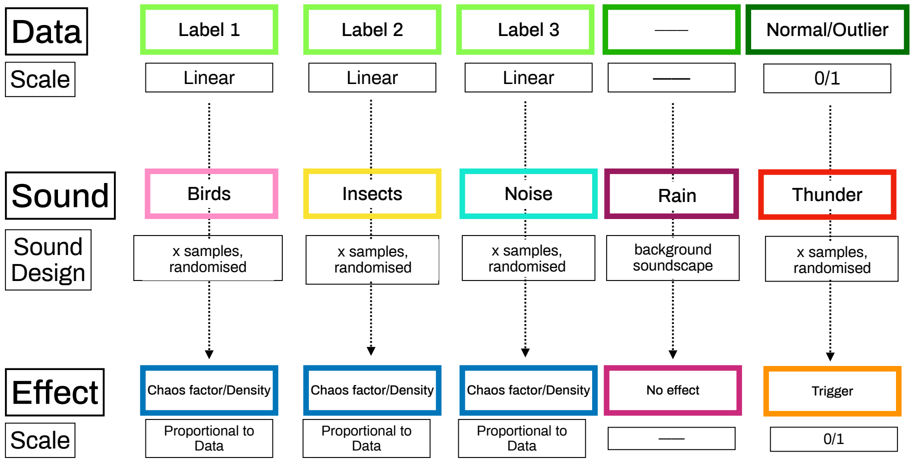


Figure 77. TECNCALIA data-to-sound mapping scheme.

Adjustments were made to the prototype interface. As shown in Figure 78, the labels were updated to show consistency with the new mapping showed in Fig.77. In particular, the rain was labeled as 'background', and the three 'labels' were named 'groups' in order to keep the prototype open for other databases. The knob for the 'background' block indicates that the parameter is not based on density but is a simple background in which different samples are played randomly.

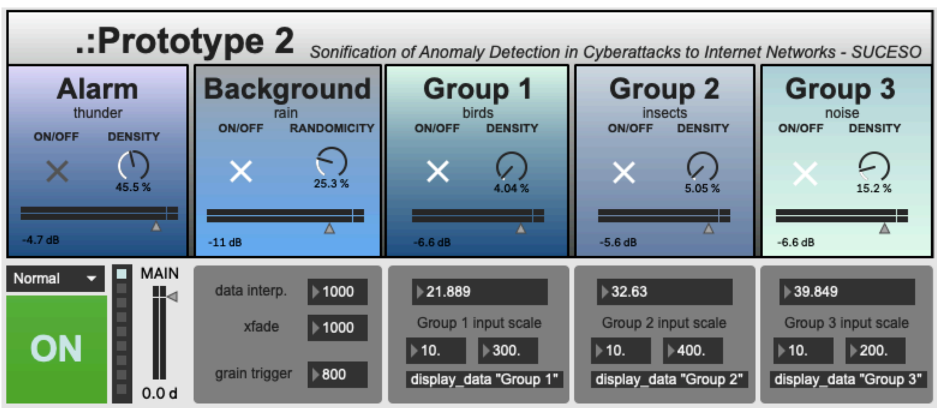


Figure 78. Version 2 of the prototype adapted for the TECNALIA testbed.

A demo video of the prototype with the sonification of both 'normal' and 'outlier' data can be found at: https://vimeo.com/515343467.

Experimental design

Two domain experts i.e., one expert in cybersecurity and anomaly detection and one expert in detection algorithms, were engaged in an experiment within the frame of the Tecnalia testbed. The experimental protocol followed the one developed for the first DA. I maintained a task-based quantitative testing with the goal of evaluating the users' performance in correctly detecting anomalous behavior in the network. As in the first DA, the quantitative testing was coupled with qualitative research to gather feedback and suggestions on the prototype according to specific topics such as the design of sounds, the user experience, the potentialities of a real-world application of the sonification. I had originally intended the qualitative part of the experiment to be inspired by phenomenological research, which has already been quoted as a possible scenario for research in sonification (Vickers and Barrass cit.), rather than by probes as in the first DA. The reason for this change was mainly due to the differences between the two Actions. In the first case, the experimental phase was meant to evaluate differences between four different prototypes and possibly lead to the identification of the most promising prototype. In the second DA, there was only prototype to evaluate, which also happened to be at a later stage of development than those in the first DA. A phenomenological approach would have involved on-site direct observations, collection of field notes as well as interviews with participants (Muratovski 2015). As I will describe shortly, the prolonged consequences of the COVID-19 pandemic, which extended to most of 2020 and are still ongoing at the time of writing, forced me to greatly limit the experimental phase. In terms of the qualitative research, the phenomenological approach had to be limited to the collection of semi-structured interviews conducted remotely after the quantitative testing.

Experimental protocol– Originally, the experiment was planned to take place at the premises of Tecnalia, a large research institution of the Basque Country, Spain, and partner of the SUCESO project. Tecnalia set up a laboratory where newly developed AI solutions can be installed and tested on dedicated machines which collect data from the local network. The environment of the lab simulates that of a typical Security Operations room, with multiple screens running simultaneously. Unfortunately, access to the lab was extremely limited over the course of 2020 and the beginning of 2021 due to local restrictions for the pandemic. Additionally, for the same reason most of the employees were working from home instead of traveling to Tecnalia's headquarters. Due to the time constraints of my doctoral project, a decision was made to run tests from the personal computers of the experts who agreed to take part in the testing from home. This decision had some consequences. Firstly, the qualitative research had to be limited to semi-structured interviews collected after the quantitative testing instead of field observations as originally intended. Moreover, despite original plans to run over several days in order to monitor changes in the understanding of the sonification after extended usage, the test ran for only three consecutive hours during the same day. Thirdly, fewer experts were able to take part in the experiment. The burdens of transitioning from office to home working and balancing work with family commitments (especially as restrictions in Spain are still severely impacting care facilities for children) made it really difficult, even for the most motivated participants, to have time for this experiment.

The final protocol included the following steps, which replicated those of the first DA, adapted to the current context:

- A presentation of the project and its goals and, more in general, of the overall concept of sonification was shared online with participants.
- A full description of the prototype, its objectives and functioning, together with details of the mapping strategy, the sound keys and audio examples of both regular and anomalous behavior was hosted on a dedicated webpage and also shared with participants and can be found at: https://www.saralenzi.com/testing [psw: Sonification2021].

It is worth noting that, unlike the first DA, in this case I decided to give testers a preview of the sonification in which participants could hear both regular and anomalous behaviour. This decision followed the results of the qualitative research of the first Action where participants lamented the absence of examples before the test and highlighted that, in a real situation, a training phase would be provided.

- A preliminary questionnaire modelled on the one used in the first DA was submitted to the participants, where they were asked to identify their main role and duties and assessed their knowledge of music and, more in general, sound culture.

- Quantitative test. During a three-hour test, participants were instructed to take notes on a spreadsheet in relation to two main occurrences:
- I heard an anomalous group. The participant was to take note of the time, identify the group and, similarly to the test conducted for the first DA, try to rate the anomaly level.
- I heard an attack. Likewise, the participant was to take note of the time of the attack and identify the group under attack.

As the reader might recall, the prototype interface contains visual cues on the status of the network and on the behavior of the sonification. Specifically, it is easy for the listener to check on the interface which sound is playing by looking at the on/off button and the volume display for each sound category (see Fig.78). Additionally, a user could see how the volume and the density knob visually increase with the increase of the anomaly level. Based on the amount of the increase, they could easily judge on the level of the anomaly. Participants were allowed to refer to the interface for a visual feedback on the sonification, as they would probably do in a real-world situation, and as was suggested by the experts involved in the first DA. Participants were required to note on the spreadsheet whether they referred to the visual interface before noting down their answer.

- Follow-up interview. Each participant was interviewed shortly after the conclusion of the quantitative testing. The interview was semi-structured around the same topics explored in the experimental phase of the first DA. Topics included: the user experience of the interface, the design of the sound, the choice of mapping strategy and the potentialities for a real-world application of the prototype. We also commented on the performance of the participant during the quantitative testing.

Risks and Limitations- Similarly to what I highlighted in the first DA, the main limitation of the experiment was the extremely low number of participants. Risks relate, mainly, to the impact of the participant's motivation on the results, as noticed in Chapter 4. Still, it is important to remember that all the participants are experts in the domain under study as professionals with at least 15 years of experience. This is a crucial factor in obtaining useful and valuable results, especially when the number of users involved in the experiment is so low.

Results

As explained, the current situation due to the spread of coronavirus obliged us to scale down the plans for the experimental phase. In the original approach, I had planned for experiments over one week to ten days of usage of the prototype during the participants' normal work routine. This would have offered sustained research on the potentialities for the adoption of the sonification in a real-world context and would have allowed me to obtain experimental results on the impact of extended training. As this first option was not viable, I opted for the application of the same protocol used in the first DA. On the positive side, this strategy allowed me to compare the results of both actions in order to gather insights and experimental evidence on the use of sonification for anomaly detection across different types of networks namely, digital and digital/physical networks. In the following paragraphs, I will share and comment the results of the quantitative and qualitative analysis for the experimental phase of the second DA. A comparative analysis of the results of the two actions is presented in Chapter 6.

Preliminary Questionnaire- As in the first DA, participants were asked to fill in a preliminary questionnaire designed to collect information on their role and expertise as well as their familiarity with music and, more in general, with sound culture. Participants were asked to self-assess their answer on a 1 to 5 Lickert scale.

One of the participants, a researcher in network security that develops tools to improve the cyber-security of industrial networks, reported a level 5 competence in cyber-security and anomaly detection but a level 1 competence in music and a level 1 attention paid to sound in everyday life.

The second participant, a data scientist and developer working in data analysis of connections and protocols in Internet networks, reported a score of 1 out of 5 in expertise on cyber-security and of 2 out of 5 in attention paid to sound. He also reported no competence in music.

Quantitative Analysis- Participants each dedicated about three hours to the quantitative testing of the sonification prototype. Testers were asked to listen to an excerpt of the Tecnalia database in which both regular and anomalous data (including data on cyber-attacks) were included. Each participant, at their best convenience, then started the sonification application they had previously downloaded on their computers. They were allowed to stop the application in case of need (to start it again from the same point) and asked to provide a detailed note on the reasons they had paused the experiment. In the event of the tester having to stop the experiment I would be able to find out why and also have precious insights for future improvements in the design of the sonification. Testers were instructed to note down, on a dedicated spreadsheet, the answers to the following questions:

1. Is the network behaving anomalously? If so, could you recognize:
   The group(s) that shows anomalous behavior.
   The level of anomaly that the group(s) is showing.

2. Is the network under attack? If so, could you recognize:
   The group(s) in which the attack is taking place.

The questions were meant to mirror those asked in the quantitative testing of the first DA where we asked participants to rate the anomaly level of the network, identify the number of anomalous districts and locate the anomaly in the specific districts where it was taking place.

Figure 79 show results of the first question - the identification of anomalous behavior in the groups. Given the extremely limited number of participants in the experiments, Fig. 79 shows aggregated results for all users. The reader might recall that the database we used did not contain a pre-determined level of anomaly to the sonification. The sound of thunder, which signals an incoming cyber-attack, is triggered only when the anomaly is labeled 'true' (in the Tecnalia case, 'outlier') by the algorithm. Other non-malicious anomalies will provoke an increase in the density of the sound within the affected group. It is the increased density that we asked participants to identify and rate on a 1 to 5 scale.

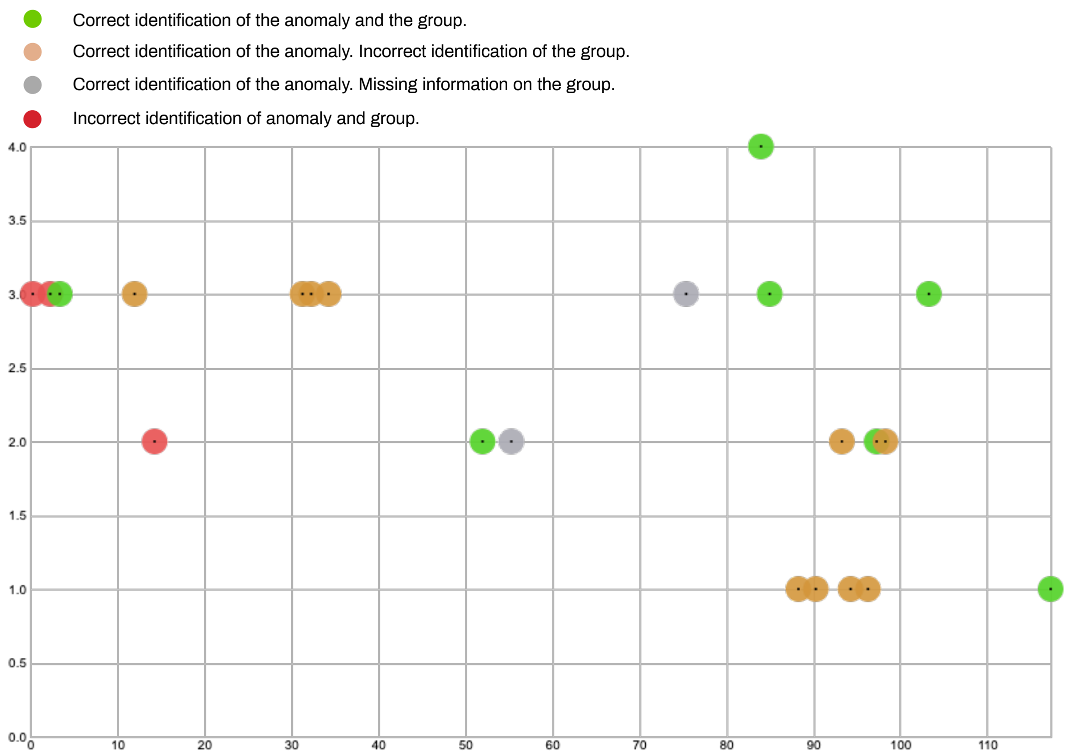


Figure 79. Quantitative analysis. Identification of anomalies and anomaly level. x-axis indicated the time, expressed in minutes, from the beginning of the test at 8AM.

Testers were asked to rate anomalies on a scale of 1 to 5, where 5 represented a cyber-attack. The maximum value attributed to anomalies for the first question, which does not include the identification of attacks, should therefore be 4. In general, and despite the participants' own feelings (as we will see in the qualitative analysis), the identification of anomalous behavior, as well as the assessment of its gravity, obtained good results (in green in Fig.79). In five cases, as shown in a light brown color in Fig.79, the anomaly was correctly identified but the anomaly level was not. To check the validity of the answers, I deducted the anomaly level from the original dataset, as in the first DA. As results in the first DA showed, a certain margin of subjectivity had to be expected since the listener is asked to rate the relative weight of the anomaly before even hearing what the maximum level of the anomaly sounds like. Despite these considerations, (indicated in red in Fig. 79) the testers only misunderstood regular behavior as anomalous on three occasions.

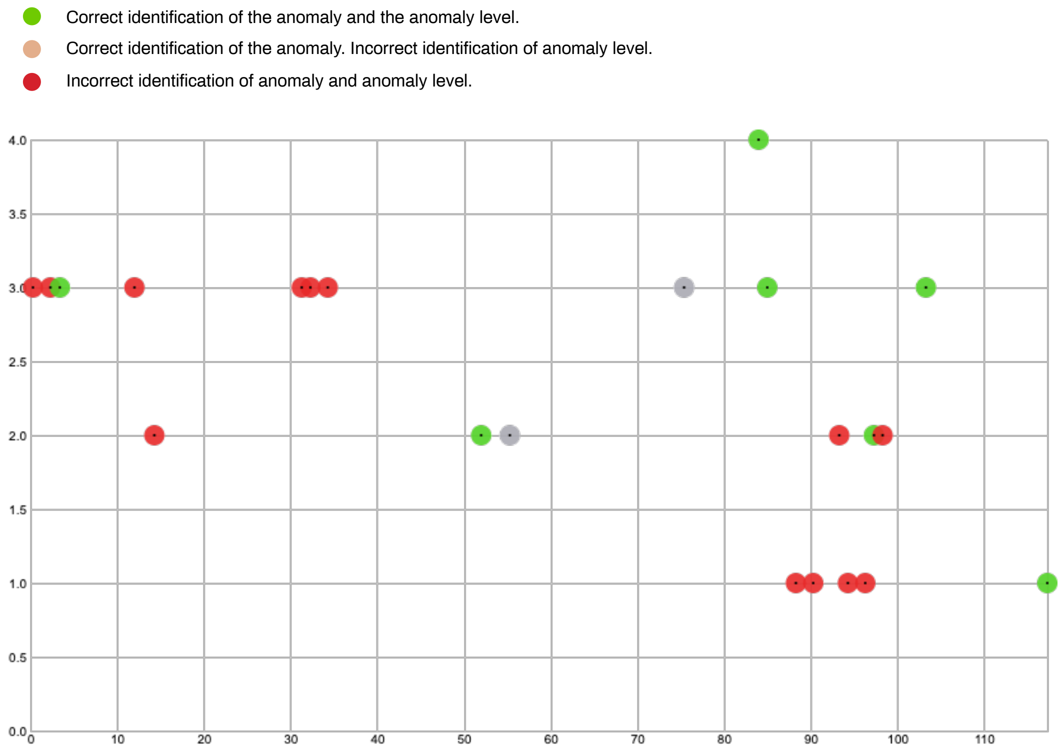


Figure 80. Quantitative analysis. Identification of anomalous groups. x-axis indicates the time, expressed in minutes, from the beginning of the test at 8AM.

Things become complicated when users are asked to identify the anomalous group among the three groups represented by birds, insects and noise.

As shown in Figure 80 (in light brown circles), participants had more difficulties in identifying the correct group after correctly identifying anomalous behavior in the network. In two cases, the anomaly was correctly identified but the users did not take note of the group due to extreme uncertainty in identifying this. Unlike the rating of the anomaly level, the identification of the group did not improve over time. As we will see, during the follow-up interview, participants highlighted some difficulties, at times, in differentiating between birds and insects.

This result offers precious insights for reviewing the design of specific sounds. A new iteration of the prototype should, for instance, increase the difference between the two sound categories to facilitate their identification by the listener.

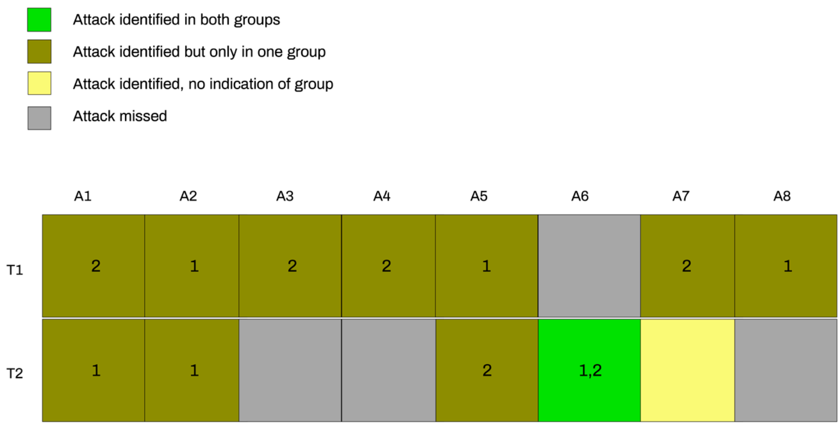


Figure 81. Quantitative analysis. Identification of attacks. In the dataset, eight attacks took place over the course of ten minutes.

In general, users performed particularly well in the identification of 'outliers'. Based on the dataset used in the experiment, attacks were concentrated during a 10 minute period towards the end of the experiment, from 10.25 to 10.35. As shown in a grey color in Figure 81, in four cases a specific outlier was missed. This might be explained by the nature of the attack, carried out with a dense sequence of several quick attacks to group 1 and 2. The short time between one incursion and the next triggers a close series of alarms (i.e., sounds of thunder) and it can be very difficult for the listener to isolate one thunderclap from another. I hypothesize that in a similar situation in the 'real world', the operator would be alerted by the sequence of thunderclaps rather than by an individual event and would therefore refer to the visualization tools for additional information. In this sense, missing a specific alarm would not be harmful for taking the correct steps in order to prevent the failure of the system but further experimental evidence is needed to corroborate this assumption. In one case, (in yellow in Fig.81) tester 2 did not indicate the group but correctly identified the attack.

Participants performed better in the identification of the group affected by the attack than in identifying a group affected by a non-malicious anomaly, however, the rapid sequence of the eight attacks poses an interesting challenge for the design of the next iteration of the prototype. Indeed, when closely looking at the dataset we can see that in this attack, both group 1 and group 2 are marked as outliers within the same minute. The reader might recall that, in the absence of an indication in the dataset of the order of arrival of data within the minimum timeframe of one minute, we decided to play the sounds in an artificial sequence. In the sequence, the frequency of the sounds during that minute depends on the amount of data received for each group, within the same minute. During at attack, data on the attack would be prioritized and the sound of the attacked group would play for the remaining duration of the minute. When the attack occurs in two (or more) different groups within the same minute, the two sounds would superimpose at a certain point during the minute and continue until new data is received. If new data is also an outlier, another (or more than one) dense sound will start creating a polyphony of sounds which is difficult to isolate. A new version of the prototype will have to carefully consider the design of each sound for each group and their possible combinations, having in mind the scenario represented in Fig.81.

A separate note should be dedicated to the usage of the visual interface of the prototype. Tester 1 checked his answers only in two occasions whereas tester 2 (‘a more visual guy’’, as he declared during the interview) checked all his answers on the visual interface. As we can see in Fig.81, no conclusions can be drawn from the choice of one or the other strategy. Tester 1 declared he used the visual cues only for attacks 1 and 3. Results show that the cue did not help him identify the correct group. The same happened to tester 2, who checked the visual cues for all the attacks. In fact, while he was the only person to correctly attribute attack 6 to both group 1 and 2, he also missed the most attacks. Perhaps the only conclusion we can reach is that the extremely limited user base prevent us from getting solid results. Further experiments are needed for specifically evaluating the performance of the sonification tool in combination with visuals.

Qualitative Analysis- In the days after the quantitative testing, participants agreed to take part in an individual follow-up interview. The interview took place using Microsoft Teams and Cisco Webex and lasted approximately 40 minutes. Following a specific request from the participants, the language used in the interview was Spanish (the full text of the interview can be found in the Appendix). As with the analysis of the first DA, comments and feedback collected during the interview are grouped in the following topics: user experience, sound design, structure of the sonification, real-world potential and suggestions.

User experience

None of the participants encountered technical issues in the download and setup of the sonification tool. The application was exported as a standalone tool for Windows which did not require the installation of Max/MSP. Participants only had to download an executable file and follow the instructions on the experiment's webpage. In general, pre-test training was judged adequate. In particular, the examples were deemed very useful for introducing the listener to differences in the level of anomalous behavior. From a technical point of view, all participants highlighted that it was not easy to take note of the time of an anomaly while listening to the sonification in real-time, but they also recognized that this task, which was vital for the testing session, would not be needed during a real-world situation. Presumably, in a real-world situation the operator would take note of the time as a consequence or at the same time as other actions, such as checking the visual dashboard or during a subsequent forensic analysis, rather than having to note the time of each anomaly – even the smallest anomaly – on a spreadsheet as was required during testing. In a couple of occasions, both participants had to stop the testing and resume it after a few minutes. They both noticed a latency in the tool which meant the sound keeps playing even if the sonification is not receiving any new data. This was problematic if the user wanted to resume the experiment shortly after: in one case, the user was not sure whether the sound that was playing was connected to the previous data or to newly incoming data. This is a minor design issue which we will take note of for a new iteration of the prototype. One of the participants lamented the fragmentation of the interface in two windows, one for the proper sonification tool and the other dedicated to the data engine simulation. The same participant also thought that in a real-life situation, data would stream directly from the data collection platform, thus presumably removing the need for a dedicated window.

In terms of the context of usage, both participants took the test from home (due to the pandemic). The environment was silent and comfortable. All users used headphones with noise-cancelling systems. They were instructed not to change the default settings for the volume of each sound group. During the interview, they confirmed that volumes for each group and for the entire sonification were adequate.

Participants were experts in network security and data science. Both listed programming as their main working task on testing day. They both highlighted that, when highly focused on a visual interface for tasks such as coding or debugging, it was difficult to dedicate even the required minimum peripheral attention to the sonification. Sometimes they even forgot they were listening to a soundscape. In their opinion, this might lead to anomalies of an intermediate level being missed. On the other hand, they felt that the sound associated with the attacks – the thunder – was noticeable enough not to be missed. We know from quantitative analysis that some of the attacks – signaled by the thunder – were actually missed by one of the participants. In the analysis (see p.161) I attributed this mistake to a specific design aspect of the current sonification that can be easily tackled. It is undeniable, though, that the issue of attracting attention through the appropriate design of sounds in the context of alarm design and sonification for process monitoring is not new (Vickers cit.) and deserves additional thorough attention ahead of a new iteration of the prototype.

Sound design

In general, the use of a natural soundscape was well received by the participants. Both experts confirmed the efficacy of the sound of thunder in attracting their attention to serious anomalies. The background sound of the rain, which was used to inform the listener on the correct functioning of the application, was deemed quite useful and pleasant at the same time. One participant said he was surprised by the positive emotions that listening to the soundscape triggered for him, to the extent that it made his other tasks more pleasant. Words like 'pleasant', 'emotional', 'relaxing' or expressions like 'I enjoyed it' are recurrent in both interviews. The sound group 'noise' was identified as the most problematic to recognize and follow whereas in general, birds and insects – and in particular, birds – were considered easier to recognize. This is in line with the original design hypothesis of a hierarchy in the perception of soundscapes which would place at the center of attention iconic sounds such as birdsong. On the other hand, contrary to the original dataset, where the three groups (in that case, the 'fields') were hierarchically ordered, the prominence of one sound category over the others could be an issue in the current testbed, where all three groups have the same importance. Other minor aspects of sound recognition, which will have to be addressed in the next iteration of the prototype, included: one of the participants identified a particularly noisy sound in the 'birds' group and commented that sometimes he wrongly attributed it to the 'noise' group. One participant shared that he found himself particularly uneasy with some of the sounds in the 'noise' category. In particular, a deep breathing sound – part of a group of sounds emitted by a sloth – was described as 'very spooky'. Whether these emotional judgements might impact the perception of the anomaly level should be further investigated with a higher number of participants.

Structure of the sonification

In general, the mapping choices to translate data into sound were well received and well understood. The reader might recall the issue of the specific management of time by the algorithm. Both participants judged as adequate the decision of artificially scheduling the arrival of data within the minute for each group, confirming that in this specific case the choice would not harm the operator's response to threats. At the same time, they highlighted that this would not be the most appropriate solution for all datasets, therefore raising issues of scalability that a new version of the prototype should take into account. One of the participants suggested adding another effect, such as a volume increase, combined with an increase in density, to attract additional attention to the variations provoked by intermediate anomalies and augment the capability to distinguish one group from the other. In particular, potential issues were raised in the distinction of intermediate anomalies due to the usage of sounds that are 'too pleasant' which might lower attention. The issue of attention lowering in process monitoring (also connected to high volumes of false alarms) known as alarm fatigue[36] (Sendelbach and Funk 2013) is well known in sonification design. Vickers (cit., p.475) states that "Fatigue is sometimes mentioned as a potential problem associated with auditory display" but it is still not clear that "auditory displays should cause more problems in this regard than visual representations". One of the testers was optimistic that an extended usage of the sonification would result in an improved capability of distinguishing intermediate anomalies.

36. "Alarm fatigue or alert fatigue occurs when one is exposed to a large number of frequent alarms (alerts) and consequently becomes desensitized to them." (From Wikipedia, the free encyclopedia).

Real-world application of the sonification

Both testers considered the usage of sonification – and specifically this sonification, based on a natural soundscape - extremely positively for peripherally monitoring anomalies in a digital network. Both declared they would appreciate the possibility to unload the visual channel and avoid having to “stare at a dashboard where nothing happens”. In the sonification design process, we decided to keep the amount of information conveyed by sound at the minimum Level 0 of the algorithm’s analysis capability. Still, participants believed there was perhaps too much information contained in the current sonification prototype. Specifically, one of the participants, an expert in network security, affirmed that in a real-world SOC the peripheral monitoring of just the attacks (instead of intermediate anomalies) on the whole network (without an indication of the attack’s location) could be revolutionary in considerably relieving operators’ current cognitive load. In his words, in digital networks many little hiccups happen at all times: the operator does not necessarily need (or want to) keep track of all of them, all the time. Additional layers of information might be useful in other contexts such as industrial production, where the margin for error is lower. According to another participant, the ability to identify intermediate levels of anomaly could guide the operator in evaluating “how fast” he has to look at the screen. In fact, after hearing the attack on the sonification, the operator would refer to the visual tools she uses on a daily basis to retrieve additional analytical information. In this sense, participants welcomed the idea that the sonification would be used as a complement, and not as a replacement, of visualizations. When asked about current alert methods on cyber-attacks, participants mentioned text messages or even e-mails: a message is sent to the operator alerting her of an attack. Besides increasing the visual channel load, these methods are asynchronous. According to participants, one advantage of using sound would be to rely on a synchronous alert method for incoming attacks. Lastly, participants shared a viewpoint that could have relevant impact on the design of the sonification. The operator, they said, does not want to know anything about the internal operations of the anomaly detection algorithm. Information regarding the type of group where the anomaly is located and intermediate levels of anomalies seemed to one of the participants to pertain to those ‘internal operations’ of the algorithm that are irrelevant and possibly confusing in a SOC. At the same time, the usage of a sonification tool combined with the existing visualizations might help keep operators more vigilant on potential discrepancies between the two sources of information i.e., on potential mistakes of the algorithm.

Suggestions, general comments

One of the participants suggested re-programing the sonification tool in Python instead of Max/MSP for ease of integration with the most common data collection engines used in network monitoring.

Despite the disruption caused by the coronavirus pandemic, I believe that engaging real experts in the evaluation of the sonification prototype in a setting that was very close to a real-world application provided valuable insights that will be integrated in the next iteration of the prototype. In Chapter 7, I will present a comparison of results obtained from the quantitative and the qualitative experiments of both Design Actions.

Part III.
Ripresa.
Design processes, protocols and tools for data sonification. Results and discussion.

Part III.
RESULTS AND DISCUSSION.

In the third and last part of this thesis, like in a conclusion of the sonata form, the ripresa, I go back to the three research questions and share with the reader the answers I have formulated following the theoretical and the practical parts of the doctoral research.

In Chapter 7, I compare the experimental results of the two Design Actions in response to the third research question.

In Chapter 8, I present a proposal for an experimental protocol for sonification projects, modelled on sonification for anomaly detection but which I intend to be applicable to sonification projects designed as solutions for real users in a real-world context.

In Chapter 9, I draft a design process for data sonification that leverages the dichotomy between reflection and action, developed in the three main phases of exploration, creation and validation that can be found in other design practices. The sonification canvas is contextualized within this more general process and, in particular, as a tool to guide reflection in the 'creation' phase.

CHAPTER 6 —

Data Sonification for anomaly detection: summary of experimental findings.

Soundtrack: *Sandro Perri, Time (You got me).*

The third research question of my doctoral path was "Can data sonification represent the complexity of digital and digital-physical systems in order to help expert users detect and prevent anomalous behavior? Can it have an impact on the daily activity of decision-makers?". It was mainly a serendipitous chain of events that led to me focusing the practical part of the research in the past three years on data sonification for anomaly detection, as this was a field I had never come across before . As I explained in Chapter 4, the first Design Action was motivated by the proposal of a collaboration by the Engineering Pillar of Singapore University of Technology and Design to explore the use of sonification for the monitoring of cyber-threats to water infrastructure. Following this first experience, during my time with the Spanish company Ibermática as part of my second-year internship, I had the opportunity to apply the findings of this first project to the real-time monitoring of digital networks for the purpose of cyber-security.

I called these two interventions grounding in practice the theoretical, overarching investigation on the potentialities of a design-driven approach to data sonification, 'design actions' (DA). The two Actions were developed following a similar process (which will be covered in Chapter 9) and were both validated through an experimental phase that applied the same protocol (with minor differences, due to the differences in the nature of the prototypes of the Actions). The development of an experimental protocol will be covered in Chapter 8. This chapter is dedicated to the summary and comparison of the results from the experiments carried out with the two DAs. Although I am aware these experimental results cannot be generalized, I still believe they can add to existing literature in the field of the design of auditory alarms and data sonification for process monitoring.

Vickers (cit., p.456) distinguishes three types of sonification for process monitoring: direct, in which the information to be monitored occupies the center of the attention of the operator; peripheral, when the attention is focused on other tasks while the operator indirectly monitors other required information; serendipitous, when the operator peripherally monitors useful, but not required, information. The case of the two DAs is the second, in which data sonification is used to peripherally monitor essential information on the status of the network. Specifically, in the first DA, four prototypes were developed to support expert users (operators of the security center of the water plant of a medium-sized city) in evaluating every hour the status of the network. In particular, operators need to be able to evaluate whether the system is behaving regularly or anomalously. In the event of an anomaly, they should be able to judge on the severity of the anomaly in order to take an appropriate decision on what actions are required. The location of the anomaly within the network's districts is also a relevant information which would facilitate the retrieval of analytical data in the existing visualization tools. In the first DA, a prototype was designed to support expert users (in this case, operators of the SOC of a medium-sized Internet network) to peripherally monitor in real-time the level of anomaly of the network. In particular, the information received from the sonification included: the presence of cyber-attacks, the level of anomaly in the event of non-regular behavior not identified as an attack and the location of the anomaly within the network. The three categories of information are, according to Guillarme (cit.), the main layers of information that an auditory alarm should convey. These three layers are what we wanted to validate through the quantitative testing of the experimental phase, for both DAs:

Layer 1. Identification of cyber-attacks- In general, participants in both experiments (for DA 1 and 2) performed well in the identification of cyber-attacks. It should be noted that while in the second DA we had a clear indication of the attack in the dataset (sonified with the sound of thunder), that was not the case in the first DA, in which the user had to identify the attack from the level of anomaly in the sonification. In both cases, both the quantitative testing and the qualitative analysis confirmed that the identification of attacks is a task that users can accomplish through peripherally monitoring a data sonification system.

Differences in the sonification system, due to the peculiarity of the network, such as the above-mentioned indication of a 'true' attack in the dataset, but also the different frequency of the sonification (real-time in case of the second DA, every hour in the case of the first DA) did not seem to have an impact on the result. For this first layer it is difficult, with the limited number of cases at my disposal, to say how much the mapping strategy is responsible for the performance. In general, in the case of the first DA, the strategy based on 'repetition' seemed to obtain better results. The second DA used a very similar strategy (based on the increase of density over time), thus corroborating the hypothesis that the increase in the repetition of sound events over a given time could be a successful strategy for the identification of the attacks.

Layer 2. Identification of anomaly level- The task of assigning a relative level of gravity to anomalies seems to be more complicated. In general, participants in both experiments performed better than they thought. The quantitative testing shows that most of the testers could associate an intermediate anomaly level, a task they admitted finding particularly challenging during the follow-up interview. Again, a mapping strategy based on the repetition/accumulation of sound over a given time seems to be promising, especially when it uses non-tuned sounds such as concrete sounds (in the first DA) or natural sounds (in the second DA).

Layer 3. Location of the anomaly in the network- The location of the anomaly in the network is where participants encountered the highest level of difficulty. In general, users managed to identify the number of areas of the network involved in the attack, but they found it extremely difficult to distinguish one area from another. This difficulty is confirmed by the results of the quantitative testing. According to Guillarme (cit.), the location of the origin of the attack in the network is a desirable but not necessary feature of auditory alarms. In the second DA, participants questioned the relevance of this layer of information which, according to one expert, is related to the internal functioning of the algorithm rather than the physical organization of the network. In the case of physical networks, such as the one under study in the first DA, spatial information could be extremely relevant. This would signify a potential difference in the design of monitoring systems for digital/physical or fully digital networks, but results are far from being conclusive. More experimental evidence is needed to assess the impact of location identification on the performance of the operator.

The reader might recall that the results of the qualitative analysis for both Actions were grouped under common topics such as user-experience, sound design, structure of the sonification and potentiality for a real-world application. We can therefore compare results of the two DAs.

Role of sound design- The importance of the design of sounds used in the sonification was confirmed by the participants of both DAs. In terms of appreciation of the design, results were not conclusive for the first DA, in which listeners were confronted with both tuned and non-tuned sound types and feelings towards the two types of sound material were mixed. It is worth noting, however, that participants tended to attribute a meaning (based on personal memories and preferences) to the non-tuned sounds, which they believed increase memorability and recognition. In the case of the second DA, participants judged the choice of natural sounds extremely positively, using words such as relaxing, pleasant, emotional. Interestingly, participants in both Actions associated feelings of dislike to the same specific sounds, which, according to them 'made them feel uneasy'. On the other hand, some sounds seemed to be particularly appreciated. Results on how much feelings of like or dislike of sounds influence the performance of users are certainly not conclusive. Interestingly, Walker and Kramer (1996, cit.) share a similar impression when they note that "We were surprised to see that the ‘Bad’ ensemble actually led to the fastest performance. The supposedly ‘Intuitive’ and ‘Okay’ ensembles led to the poorest performance, overall, while the ‘Random’ ensemble led to the best performance”. Nevertheless, results from the qualitative analysis seem to indicate that positive feelings about the design of sounds in the sonification would greatly influence the motivation in adopting the sonification in a real-world context. I believe that these results – although they are far from being conclusive – highlight the importance of the role of aesthetics in sonification design and suggest this is not merely a cosmetic operation but rather a fundamental aspect of the relationship with the user, as highlighted on several occasions by members of the sonification community (Vickers and Barrass cit.; Furlong and Roddy cit.; Barrass 2012 cit.).

Structure of the sonification- In general, the frequency of the sonification, i.e., real-time in the case of the second DA and every hour in the case of the first DA, was positively welcomed by the participants. Both groups confirmed that this aspect is highly dependent on the characteristic of the network under study and in particular, on the time that an attack would need to provoke the collapse of the system. These considerations highlight the importance of the correct definition of all aspects and constraints of the specific use case to which the sonification is dedicated. In terms of the specific mapping strategy, as mentioned in the previous paragraphs, results from the quantitative testing seem to indicate that the use of repetition/accumulation of sounds representative of the anomalous group over a given time is promising. This result deserves to be tested across different use cases of sonification for anomaly detection in order to explore its potentiality. New testing should also address the issue raised by participants in the first DA, of the lack of a scale of comparison for the number of repetitions as well as the influence of different types of sounds on user performance. The role of visual cues, introduced in the second DA, is not clear. Participants recognized the importance of this during the training phase but were doubtful whether it could be of support in a real-world context, where other visualization tools, which are already well known to the operator, would be more important.

User experience- The experiments confirmed the use of headphones in the specific context of sonification for process monitoring, which gives precious indications for the sonification design. In particular, the usage of spatialization of sounds for example in relation to the localization of the anomaly (as in Edward and Ville 2003; Iber et al. 2020), could be explored in future iterations of the prototypes. Other issues, such as the need to adjust the volume in the case of conflicting sound events (such as incoming phone calls) and the need to hear the sound when the operator is distant from her desk could be addressed by design, for instance by including an automatic volume correction or developing a portable application for the sonification.

Real-world application- One of the main premises of this thesis is that a designerly approach to sonification can enhance its possibilities of becoming part of the data representation solutions adopted in real-world contexts. It was of the highest importance, consequently, to gather feedback on the potentialities of the DA prototypes as real-world applications. Participants of both Actions were extremely optimistic that sonification can become a widely adopted method of monitoring networks' anomalies, thus liberating an overloaded visual channel from the need to 'stare at a screen where nothing happens'. This confirms one of the main assumptions of both projects.

As mentioned, the quality of the sound design was quoted as one of the main motivations for the adoption of the sonification tool during normal work routine. Participants in the second DA confirmed that the use of natural soundscapes is promising as it can seamlessly integrate with the work acoustic environment. Both groups highlighted the importance of training in real-world contexts in order to refine the capability of understanding the information conveyed by the sonification, and in particular the level of anomaly. In this sense, users of the second DA appreciated the visual cues contained in the prototype interface. On the other hand, the utility of such cues in a real-world context was debated, as operators would prefer to use existing visualizations they are already familiar with. Both groups agreed that the sonification should complement, and not replace, existing visual dashboards. This last point suggests an urgent need for experiments to assess the performance of the sonification prototypes in combination with existing visualizations.

In conclusion, experimental results encourage the use of sonification for the monitoring of anomalies in the context of digital and digital/physical network. The strategies we adopted, in particular the usage of repetition/accumulation of sound over time to signal an increase in the anomaly level and the usage of a sonic cue (thunder) to attract the attention on ongoing attacks, seem to be valid but need further experimental evidence. The efficacy of the use of sound to locate the anomaly within the network is not clear and should be further tested. The attention paid to the design of sounds by users is key and, as we will see in the following chapters, is a key indicator that sonification prototypes should be tested with real users in a real context.

CHAPTER 7 —

Defining an experimental protocol for data sonification.

Soundtrack: *Brian Eno, Music for Airports*

In the first chapter of this work supported by the literature, I posited that the lack of a culture of experimental validation in the sonification world negatively impacts the potential of sonification to become a widely recognized and adopted means of data representation. The recognition of this gap led to the formulation of the second research question 'How can we evaluate whether a sonification is efficient and engaging for a user? Can we frame a design methodology to approach data sonification projects from prototyping to testing?'

During my doctoral project, I had the opportunity to research and validate, through the two Design Actions, an experimental protocol for the evaluation of sonification projects. In particular, the experimental design focused on the evaluation of sonification prototypes which were meant to support domain expert in anomaly detection tasks. Due to unforeseen circumstances, the protocol developed for the first DA was also applied to the assessment of the second DA. Based on these two cases, I distilled the proposal – which I present in this chapter - of an experimental protocol for data sonification projects that target expert users with the goal of supporting them in their daily tasks, in combination with existing visualization tools.

In general, both DAs were validated using a combination of quantitative and qualitative research. In the case of the first DA, six domain experts were engaged in an experiment which took place over four days, eight hours a day. Participants were presented with four different early-stage sonification prototypes, which they used to perform several tasks in the context of anomaly detection in a water distribution network. A key requirement of the experiment was that participants use the sonification during a normal working day, in their usual workplace and while carrying out routine tasks. Participants were selected based on their familiarity with the domains involved in the study: water infrastructure engineering and cyber-security. Before the experiment, all participants were asked to complete a questionnaire to evaluate their prior knowledge of music and, more in general, their sound culture, as well as to describe their background and their role. Information on the topic of sonification, the context of the project and the experimental protocol was collected on a purposely designed web page and shared with each participant during a training session organized via Skype (or equivalent) followed by a Q&A session in which each participant could ask questions or clarify doubts.

During the quantitative testing, participants were asked to answer several questions in order to evaluate their performance in relation to the main goal of the sonification i.e., supporting expert users in anomaly detection. In particular, they were asked to evaluate:

- The status of the system (regular or anomalous).
- The gravity of the anomaly.
- The occurrence of a cyber-attack.
- The location of anomalies and attacks along the network.

In the days after the quantitative testing, a one hour-long follow-up interview was conducted (remotely, via Skype or equivalent) with each expert. The interview was meant to explore, in a semi-structured form, specific topics relating to the design of the sonification prototypes. It was meant to help us better contextualize the results of the quantitative testing and gather insights and suggestions for a new iteration of the prototypes. The topic explored during the interview roughly covered:

- The user experience i.e., the overall ease of use, technical performance and experience of the participants in their first encounter with the sonification. This included gathering information on the sound diffusion system that was used, the characteristics of the environment where the test was conducted, the actions and tasks carried out by the participant during testing, the characteristics of the environment in which the testing took place.
- The design of the sonification. We tried to gather as much feedback as possible on topics such as mapping strategies, choices relating to the types of sound, including their aesthetic quality and functionality, and the appropriateness of the frequency and duration of the sonification during a normal working day.
- The pre-test training: we asked for feedback on the quantity and quality of the training provided before the test, both in general about principles of data sonification and about the specifications of the prototypes.
- Finally, we shared the results of the quantitative testing and commented on it with each participant.

This protocol, as described in chapter 4, was largely inspired by the use of probes in design research as a method to collect information on a design artifact in a real environment with an additional constraint imposed by the specific goals of the sonification: that participants in the experiment would be domain experts i.e., that they would be familiar with the tasks required by the quantitative testing and that they would be able to share insightful feedback on the adoption of the sonification in a real-world context.

For the second DA, I tried, despite the restrictions and difficulties resulting from the COVID-19 pandemic, to replicate the same experimental protocol. Two domain experts, in this case professionals in the fields of cyber-security, digital networks or anomaly detection, were engaged to test one late stage sonification prototype, in a setting that was as close as possible to a real-world setting, during a normal working day, with the goal of monitoring anomalies and detecting cyber-attacks to an Internet network. A preliminary questionnaire, which replicated the one used for the first DA, was submitted by the participants. The quantitative testing ran for one day over three consecutive hours. Participants were asked to answer questions analogous to those asked during the first DA testing. Questions aimed at evaluating the participants' performance in the identification of anomalous behavior and attacks to the network, rating of the level of anomalies and area of the network (in this case, the group) in which the anomaly had occurred. A follow-up interview covered the same areas as the first DA i.e., mainly, sound design of the prototype, overall user experience, adequacy of prior training. Results of the quantitative testing were also shared and commented. Prior to the pandemic, the inspiration for this second experiment was phenomenological research. I had planned to engage the participants in a prolonged use of the prototype over several working days, in the hope of evaluating the impact of extended usage on user performance and the potentialities of the prototype as a real-world solution. As described in Chapter 5, limitations resulting from the pandemic forced me to restrict the scope of the experiment which, nonetheless, still kept the main constraints of a real-world setting (at least temporarily, in light of COVID) and the evaluation of the sonification by domain experts.

In response to the second research question, I propose an experimental protocol for the evaluation, in a real-world setting by domain experts, of sonification applications designed for the purpose of anomaly detection. The protocol, which is illustrated in Figure 82, has two main constraints (the real-world setting and the engagement of domain experts) and an obvious limitation. It is distilled based on two design actions during an unusual and unique situation by an undeniably small group of participants.

Nonetheless, I believe that the combination of quantitative testing and qualitative research offers the designer a good perspective on the real potentialities of the sonification application at a prototyping stage. The results of the experiment can (and in the case of the two DAs, do) offer valuable insights both in terms of the efficacy and efficiency of the application and its potentialities as a real-world product. In terms of quantitative research, the setup of the testing allows the designer to gather precious evidence on the main requirements that sonification used as an intelligent auditory alarm has to satisfy, according to

Guillarme (cit.). Notably, it should indicate how serious a failure is, what caused it and the location of the fault. In terms of qualitative research, during the interview participants have the opportunity to share their informed opinion, as domain experts, on the most promising characteristics of the sonification as well as potential issues in real-world applications. In the case of the two DAs, experts also gave useful and interesting advice on possible evolutions of the prototypes.

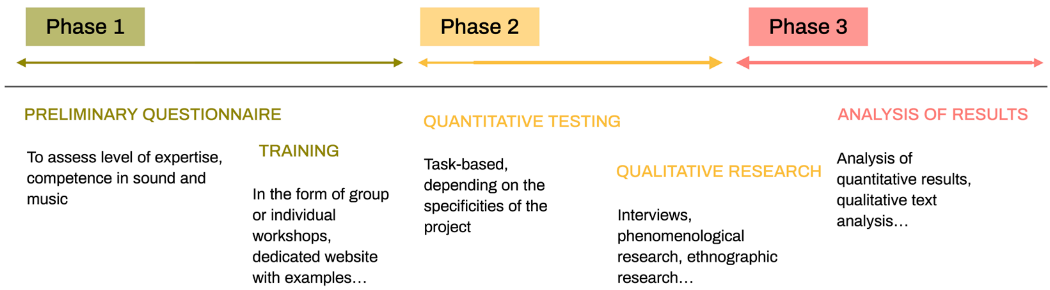


Figure 82. Experimental protocol for sonification for anomaly detection.

As shown in Fig.82, the protocol is organized in four steps grouped in three phases. A preliminary phase, which includes the first two steps and is rolled out in preparation of the two main phases, the quantitative and the qualitative experiments and the analysis of the results:

Phase 1 is composed of:

1. A training workshop, in which information about the practice of sonification, the specific project, the application as well as audio examples and details of the experimental protocol are shared. During our experiment, we found that collecting all this varied information on a web page to share with participants beforehand, seems to be a good solution. A follow-up one-to-one session should be arranged to recap the general information and specific instructions of the experiment, and clarify any doubts.

2. A preliminary questionnaire, to gather demographic information and evaluate the level of expertise of the participants in both the domain under study and in terms of music and sound.

In *Phase 2*, experts take part in:

3. A quantitative test, that aims to assess their performance in using sonification to detect anomalous behavior. As already

noted, it seems from our experiments that the criteria presented by Guillarme (cit.) are appropriate. Based on these criteria, the quantitative testing should evaluate the efficacy of the sonification in facilitating the detection of anomalies, the correct judgement of the severity, and the correct location of the anomaly in the network.

4. A subsequent qualitative assessment, based on semi-structured interviews with participants that aim to assess specific areas of the sonification design such as the mapping strategy, the types of sounds, the user experience and the potentialities for real-world application, as well as comment on the results of the quantitative testing. I believe that other methods such as field notes and observations could be integrated. The specific qualitative methods (for instance, a probe-based rather than a phenomenological or ethnographic method) should be chosen based on considerations such as the status of the prototype (early or late-stage, one or more versions), the specific goals of the sonification (monitoring of digital or digital/physical networks), the context of usage (private or public company, strategic infrastructure, etc.), and so on.

Phase 3 is dedicated to the analysis of both quantitative and qualitative material collected during phase 2.

5. Quantitative analysis will depend on the characteristics of the sonification, the specific tasks to be performed during the testing and the number of participants. Qualitative analysis will include processes borrowed from qualitative text analysis protocols.

Given the relative short duration of both experiments and the small number of participants, it is too early to say whether the experimental protocol I propose can be effectively replicated and generalized, within and possibly beyond the specific use case of sonification for anomaly detection. Nonetheless, and more in general, I consider the attention paid to the engagement of real potential users in a real-world setting a promising approach for the design of sonifications that can have an impact on our daily relationship with data.

CHAPTER 8 —

Defining a design process for data sonification

Soundtrack: *The Strokes, The Adults are Talking*

The first research question (What is the role of the (sound) designer in data sonification? Can a designerly approach to sonification make the difference in creating better - more efficient and engaging - representations of data?) addressed the potential role of design as a discipline to provide a framework for the conceptualization, design and implementation of sonifications that could have an impact in the real-world. Through the course of this thesis, the role of design has been investigated in two ways. On one side, as a sound designer, I engaged in the design and production of a series of sonification prototypes in the frame of the two Design Actions. This provided the opportunity to apply data sonification to real-world problems and subsequently reflect on this practice in order to distill guidelines for the application of design methods to data sonification projects. On the other side, through interviews with sonification experts and workshops with communication designers, I conducted an investigation into the development of specific design tools to support designers who are not familiar with data sonification in the integration of the auditory sensory modality in the representation of data. The results of both these approaches are illustrated in the following paragraphs.

A design-driven work process for data sonification - In Chapter 1 I explored, among other things, how practitioners of specific areas of sound design, notably sound design for film, sonic interaction design and audio branding, tend to apply a rather standard 'design thinking' work process to the development of their projects. This general method roughly covers a preliminary research phase, a central creative phase and a subsequent phase in which prototypes or final products are shared with users or clients to gather feedback prior to the final implementation. The path I followed in both DAs seems to inductively confirm that this general design approach can work for data sonification projects. From the specific steps and phases undertaken during the design and development of the DAs (illustrated in Chapter 4 and 5) I distilled a work process for the design of sonifications which is shown in Figure 82. This process is meant to support the sonification designer through three main phases – exploration, creation and implementation. As an overarching approach, each phase is characterized by an oscillation between reflection and action. The use of this dichotomy is inspired by Donald Schön's seminal 'The reflective practitioner'. In particular, I find these two terms more appropriate than the originally chosen terms 'theoretical' and 'practical' because, with Schön, "When someone reflects-in-action, he becomes a researcher in the practice context. He is not dependent on the categories of established theory and technique, but constructs a new theory of the unique case." (p.68). In the reflection-in-action process, theory and practice are interactively defined as the designer continuously frames (and re-frames) the problem as it unfolds. This oscillation and mutual feedback process of action and reflection is described in Fig.83. Specific tools and deliverables are attached for both the Action and the Reflection poles, for each phase. The definition of each of these elements – the real, specific contents of an otherwise standard design process – is the real contribution of my work. Each step is meant to guide both the sonification expert and the communication designer through the creation of a new project. On the one hand, the sonification expert would benefit from an integrated process which applies the lenses of design, a structured discipline used to tackle complex, real world problems, to the specificities of data sonification. On the other hand, the process is meant to help the communication designer integrate sound as an alternative sensory modality to represent and communicate information. With this goal, purposely developed tools, such as the sonification canvas, which target the non-expert in sonification, are included in an otherwise familiar design methodology.

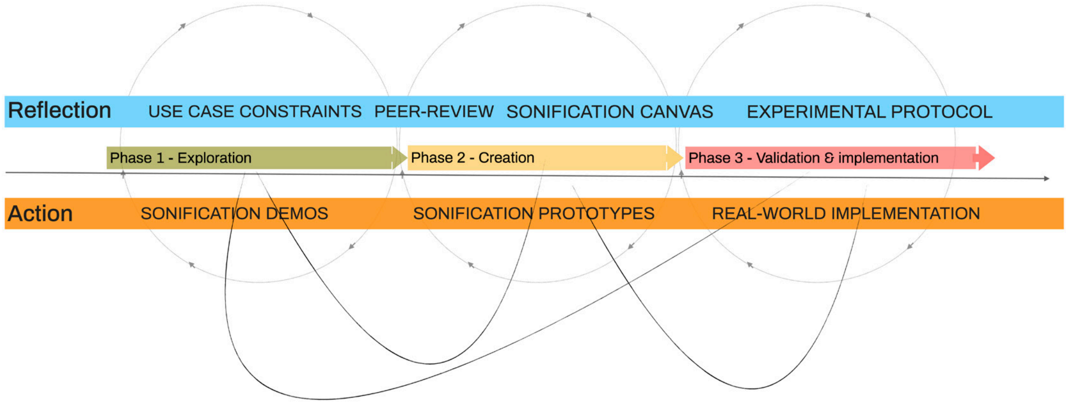


Figure 83. Data sonification design process.

Phase 1. Exploration

Reflection- During the exploratory phase of a new sonification project, the designer conducts preliminary research on the topic of the sonification, the real-world phenomenon and the specific dataset. Any other pre-defined constraint, such as, for instance, technical aspects of the sound diffusion system, minimum requirements on the information that the sonification should convey (does it have to improve the performance of specific tasks?), characteristics of the target users (are they visually impaired, do they belong to a specific cultural group, do they have a specific expertise or a specific opinion on the phenomenon under study? and so on), characteristics of the context of usage (is it a site specific installation, a web experience?) or the relationship with other representation modalities (is the sonification associated to an existing visualization?) should be taken into account and investigated.

The deliverables attached to the 'reflective' viewpoint of the first phase would be supporting documentation including:

1. Definition of the use case
2. Definition of the design constraints

Action- While the preliminary research is ongoing, the production of one or several sound demos is recommended. These demos are meant to freely explore possible mapping strategies and types of sound without constraining the design to the specificities identified in the 'reflection' part of Phase 1. From this exploratory practice, the designer will gather as much insight as possible on potential options that can be later integrated during the prototyping phase. Demos can be structured as musical sonifications, artworks, and can combine visual elements, explore sound material of any type or be used to test initial requests by a client.

Deliverables of this phase include:

1. A variable number of sound demos with no specific constraint of duration, format or genre.

2. Demos should be shared with team members, partners of the project or clients during closed-door expert sessions in order to gather early feedback.

Phase 2. Creation.

Reflection- In this phase, the sonification designer takes core decisions on the mapping strategy, the type of sounds and the technical requirements of the sonification. As a guide for the decision-taking process, I suggest the usage of the sonification design canvas which has been thoroughly illustrated in Chapter 2 and will be presented again in the following section. Despite specifically targeting communication designers, the sonification canvas can be a valid guide to support the expert in sonification in structuring the design process.

Deliverables of this phase include:

1. Different versions of the sonification canvas, that should be always kept up to date during the various stages of the project.

2. Other supporting material such as prototypes sketching, co-creation visualizations, music scores, software mockups and so on.

Action- One or more prototypes are conceptualized and designed accordingly. Depending on the context, prototypes can be early-stage 'probes' to be used in preliminary design explorations, as in the first DA, or late-stage fully functioning digital tools which have already undergone a series of design iterations, as in the second DA. The creation process is iterative: the reflections on the work process and on the choices made by the designer and the actions taken for the concrete design of the prototypes inform each other in a feedback process that continues during Phase 3.

Deliverables of this phase include:

1. One or more prototypes, in different versions and at different stages of development, depending on the project.

Phase 3. Validation and implementation.

Reflection- The prototypes are evaluated with an appropriate methodology. In Chapter 8, I propose a specific protocol for the evaluation of data sonification and in particular, of sonifications for the detection of anomalous behavior. The protocol is a combination of quantitative and qualitative research and aims to evaluate early stage and late stage sonification prototypes with real users in a real-world context. Results of the experiments are analyzed and used to inform subsequent iterations of the prototypes.

Deliverables of this phase include:

1. A document detailing the experiment procedure and contents.
2. A repository of preliminary information to be shared with participants.
3. Details of questions, testing and other tools for the collection of information from participants.
4. A document presenting the results of the analysis, including, if applicable, visualization of the experimental data.

Action- The experimental cycle informs updates and revisions of the sonification prototypes until the final product is finalized for implementation in the real-world. At this stage of development of the field of sonification, implementation as real-world product is an ambition to which the sonification designer should aim, rather than a common achievement. I include it in Fig.83 as the ultimate goal to which designers should aspire for sonification to transition to a real medium of communication that can impact our relationship with data.

Deliverables of this phase include:

1. Sonification as a real-world tool or experience.

The sonification design process I present in this chapter does not have the ambition to radically change how authors of sonifications conduct their projects and even less so, innovate design processes per se. It is proposed as a simple guide, inductively distilled from the process of design and development of the two Design Actions. It targets the sonification expert (rather than the design expert) who may not be familiar with design processes and methods. I believe that the process outlined in Fig.83, that largely mirrors a typical design approach that is already in use in other fields of sound design (see Chapter 1), can be successfully adopted by authors approaching sonification as a real-world application[37]. I consider the adoption of a common work methodology by sonification designers as a much-needed step to allow the sonification community to share, compare and replicate, on a common basis, the results from a variety of sonification projects. This will in turn, over time, help create a corpus of research and practices in which sonification can ground its roots to sustain its transition from a niche field to a widespread design practice based on common principles.

37.
It would be less relevant, for instance, for researchers working in basic research in sonification.

Tools. The sonification canvas- The sonification canvas was developed from the categories that emerged from interviews with sonification experts and validated through two workshops with communication designers (see Chapter 3). Unlike the sonification design work process, which is meant as a support for experts in sonification and as a contribution for the advancement of the field of sonification, the canvas is meant to facilitate the integration of data sonification in the daily practice of the communication designer. For the reader's ease of consultation, Figure 84 presents the sonification canvas in its latest iteration. As I will explain shortly in the final remarks of this thesis, I plan to dedicate further research to the application and validation of the canvas as well as to the development of other design tools for communication designers who want to use sound as a novel means of representing data.

Figure 84. The sonification canvas.

Conclusions and future developments.

I started this work with an admission: that my investigation on data sonification was born out of skepticism (or even disbelief) that sound could really be a 'good' (effective, efficient and engaging) way to represent data and contribute to the needs of an increasingly data-intense society. Without possessing a formal training in design disciplines, but with more than ten years experience in the practice of sound design, I set out to evaluate to what extent, and how, a 'reflective practitioner' (Schön, cit.) could enrich the debate on the transition of data sonification from a specialized means of scientific analysis to a cross-disciplinary medium for the representation and communication of complex phenomena for a variety of purposes and users, in a variety of contexts. The search for a design-driven approach to data sonification led me to reflect, on the one hand, on the tools and processes involved in using sound for data representation. In Chapter 1, I explored how existing sound design practices (sound design for film, sonic interaction design, sound for gaming, sound branding) define their design space and reflect on the process of creation and validation. In Chapter 2, I turned my attention back to things, entering into a dialogue with authors of sonifications with the goal of gathering insights on their practices, expectations and vision for sonification. From this exploration, I condensed a tool – the sonification canvas – that is meant to support communication designers in integrating sound into their representations of data. Chapter 4 and 5 are dedicated to a reflection on two 'design actions' through which I had the opportunity to explore and ground my approach to sonification in the solution of a real-world problem: the detection of anomalies in digital and digital/physical networks.

Through the actions, I tried to delimit complex, wicked problems (Buchanan cit.) related to complex, evolving real-world phenomena, framing them around real users, in a real context. In Chapter 8, I formalized a data sonification design process that aims to support the sonification designer in intentionally designing sonifications with a specific communicative goal for specific users, in a specific context. I regard the formalization of a design process for data sonification, as well as the definition of specific tools in support of this process (such as the sonification canvas), as a much-needed step to ground the relevance of sonification in a shared corpus of research and practice. The field of data sonification does not need a 'killer app' (Supper cit.) but it can greatly increase its impact in the real world through the progressive definition, case after case, of a common design space.

I learnt from my graduate and post-graduate studies in philosophy and history of science that a science is built first and foremost through the validation of peers (Kuhn 1962; Shapin and Schaffer 1985). In Chapter 7, I propose an experimental protocol for the evaluation of sonification projects that combines quantitative and qualitative research and is modeled on sonification for anomaly detection. I believe that an experimental approach, tackled from an early design phase, is key to the development of successful sonifications through the design of subsequent iterations that are informed by an ongoing conversation with real users. It also provides a community – in this case, the sonification community – with a common space and with specific instruments and methods to define the borders of its practice. Lastly, experimental results from the two design actions, presented in Chapter 6, seem to confirm that data sonification can be used to build valuable tools for the early detection of anomalies in digital and digital-physical networks.

This work proposes an experimental protocol, a design process, a specific tool and a prototype for the real-time detection of anomalies in digital networks. To sustain these four achievements – that are far from being conclusive – I plan:

- To promote the usage of the sonification canvas by communication and sonification designers through workshops and other dissemination activities. This will help evaluate its potentiality and iterate its design, but also reach a critical mass of projects that could help me gather additional insights on how the world of sonification is evolving.

- To foster the application of the experimental protocol and of the design process, first and foremost by adopting them in my own future data sonification projects, in order to evaluate their adaptability to other contexts and users as well as to gather precious experimental results.

- Both activities will be run from the newly launched sonification. design web repository (Lenzi et al. op. cit.) which has the ambition to become a reference point for designers and sonification scholars and practitioners.

- Lastly, I plan to continue the development of the second DA prototype. A new version of the prototype, that will integrate the results of the first series of experiments, is already under development. The new prototype will be validated through a second series of experiments, potentially in a post-pandemic context. These experiments aim to evaluate the new prototype engaging a higher number of expert users and will also try to explore changes in the performance of users over a longer period of time, as originally planned. On the other hand, I envisage applying the prototype in other fields where anomaly detection plays an important role, such as healthcare, industrial processes and digital-physical networks like water and electrical infrastructure.

In the Introduction to this work, I collected and analyzed several definitions of data sonification. Now that I have reached the end of my doctoral project, I consign to the reader my - temporary - definition of data sonification as the use of sound, alone or in combination with other sensory modalities, to improve the relationship with data of specific users, in a specific context and with a specific purpose. To the communication designer, this definition will sound familiar in that it refers to users, contexts and purposes, the necessary ingredients of any communication recipe. It adds a component that is mostly unfamiliar to the design community, i.e., sound: the unexpected gluing element that can change the way we represent and communicate data. To the sonification expert, this definition might sound too broad, as it does not restrict the material we use, or the goals of our work. It also does not attempt to provide a solution to the mapping problem and other daunting issues of sonification. In fact, I propose framing these and future issues as design problems that can be better tackled if faced case by case, constraint after constraint. And, by being shared with the community, the sonification design problems of specific projects might help shape best practices that really impact the way we relate to data.

The work of the past three years aims to reach out to two worlds, an ambition that is also a wish for the future: to the sonification expert, I wished to provide instruments to tackle sonification through the lenses of design as a structured discipline that successfully confronts complex problems by providing multiple solutions that have a real impact on people's lives. And I wish to make communication designers aware of the immense potentiality of sound, a sensory modality that, whether we like it or not permeates our lives and shapes our experiences. By consciously integrating it in the way we represent and communicate data, we could improve the chance to return the real richness of complex phenomena that are behind datasets.

References

Alexander, C Silverstein, M and Ishikawa, S. (1977) A Pattern Language: Towns, Buildings, Construction. Oxford: Oxford University Press.

Anceschi, L (1998) Che cos'è la poesia? Bologna: CLUEB.

Asutay, E, Västfjäll, D, Tajadura-Jiménez, A Genell, A Bergman, P and Kleiner, M. (2012) Emoacoustics: A Study of the Psychoacoustical and Psychological Dimensions of Emotional Sound Design. In J. Audio Eng. Soc., vol. 60, no. 1/2, pp. 21-28, (2012 January.)

Louise Axon, Bushra A AlAhmadi, Jason R C Nurse, Michael Goldsmith, Sadie Creese, Data presentation in security operations centres: exploring the potential for sonification to enhance existing practice, Journal of Cybersecurity, Volume 6, Issue 1, 2020, tyaa004.

Axon, L., Happa, J., Goldsmith, M., and Creese, S., (2019) Hearing Attacks in Network Data: An Effectiveness Study in Computers & Security. 83:367-388.

Axon L, Nurse JRC, Goldsmith M, Creese S. (2017) A Formalised Approach to Designing Sonification Systems for Network-Security Monitoring. In: International Journal on Advances in Security. Vol 10., 2017:26-47.

Ballatore A., Gordon D., Boone, A P. (2018) Sonifying data uncertainty with sound dimensions. Cartography and Geographic Information Science, 46(5): 85-400.

Ballora M, Pennycook B, Ivanov PC, Glass L, Goldberger AL. (2004) Heart Rate Sonification: A New Approach to Medical Diagnosis. Leonardo, 37(1):41-46.

Bang, A. L. and Eriksen, M. A. (2014). Experiments all the way in programmatic design research. Artifact, 3(2), 4-1.

Binder T. and Redström, J. (2006) Exemplary Design Research, In Friedman, K., Love, T. and Corte-Real, E. (Eds.) Proceedings of Design Research Society Wonderground International Conference 2006.

Ballora, M., Pennycook, B. W., & Glass, L. (2000). Audification of Heart Rhythms in Csound. In The Csound Book: Perspectives in Software Synthesis, Sound Design, Signal Processing, and Programming, edited by R. Belanger. Cambridge (MA): MIT Press

Bakker, S., van den Hoven, E., & Eggen, B. (2012). Knowing by ear: leveraging human attention abilities in interaction design. Journal on Multimodal User Interfaces, 5(3-4), 197–209.

Barrass, S (2016). An annotated portfolio of research through design in acoustic sonification. Leonardo, 49(1), 72-73.

Barrass, S (2012) The aesthetic turn in sonification towards a social and cultural medium. AI & Soc 27, 177–181.

Barrass, S (2003) Sonification Design Patterns in Proceedings of the 9th International Conference on Auditory Display (ICAD2003), Boston, MA, July 7-9, 2003.

Barrass, S (1997), Auditory Information Design. PhD Thesis, Australian National University, Canberra.

Bauer, M. W. (2000). Classical Content Analysis: a Review. In Bauer, M. W., & Gaskell, G. (Eds.), Qualitative Researching With Text, Image And Sound: A Practical Handbook (pp. 131-151). London, United Kingdom: Sage Publications.

Beckerman, J and Gray, T (2014) The Sonic Boom: How Sound Transforms the Way We Think, Feel, and Buy. Boston, MA: Mariner Books.

Bishop, MJ and Sonnenschein, D (2012) Designing with sound to enhance learning: Four recommendations from the film industry, The Journal of Applied Instructional Design · Volume 2 · Issue 1, pp.5-12).

Blattner, M. Sumikava, D, Greenberg, F. “Earcons and Icons: Their Structure and Common Design Principles”. In Human-Computer Interaction, vol. 4, 1989, pp. 11-44.)

Brazil, E and Fernström, M (2009) Subjective experience methods for early conceptual design of auditory display in Proceedings of the 15th International Conference on Auditory Display (ICAD2009), Copenhagen, Denmark. May 18-22, 2009.

Bonebright, T and Flowers, J (2011) Evaluation of auditory display in Hermann, T., Hunt, A. and Neuhoff, J. G., eds. The Sonification Handbook, Chapter 6. Berlin, Germany: Logos Publishing House.

Bremner, C. and Rodgers, P. (2013). Design Without Discipline. Design Issues, 29(3), 4-13.

Briggs, J. (2016) https://www.revealnews.org/blog/a-sonic-memorial-to-the-victims-at- orlandos-pulse-nightclub/Last access October 27th, 2019.

Briones, M. A. (2018) https://alternative-narratives-vis-archive.com/case_studies/egypt-building-collapses.html/ Last access May 27th, 2020.

Bronner, K, Ringe, C and Hirt, R (2020). Audio Branding Yearbook. Baden-Baden: Nomos Verlagsgesellschaft.

Buchanan, R. (1992). Wicked Problems in Design Thinking. Design Issues, 8(2), 5-21.

Cairo, A. (2019) Sonifying Data with TwoTone, https://twotone.io/about/ . Last accessed 12th Jan 2021.

Chion, M. (1994), The Audiovision: Sound on Screen. New York, U.S.: Columbia University Press.

Chion, M. (2015). Sound: An acoulogical treatise. Durham, NC: Duke University Press.

Collins, K. (2008) Game Sound: An Introduction to the History, Theory, and Practice of Video Game Music and Sound Design. The MIT Press.

Cory, M. (2015) https://www.revealnews.org/article/listen-to-the-music-of-seismic- activity-in-oklahoma//Last access October 27th, 2019.

Cross, N. (1982) Designerly way of knowing, Design Studies, Volume 3, Issue 4, October 1982, Pages 221-227.

Dayé, C. and de Campo, A. (2006.) Sounds sequential: sonification in the social sciences, Interdisciplinary Science Reviews, 31(4): 349-364.

de Campo (2007) A ata sonification design space map in Proceedings of the 2nd International Workshop on Interactive Sonification, York, UK February 3, 2007.

de Campo, Frauenberger and Höldrich (2004) Desgning a Generalized Sonification Environment in Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display, Sydney, Australia, July 6-9, 2004.

Diaz Merced, W. L. (2013) Sound for the exploration of space physics data, PhD thesis, University of Glasgow.

Debashi, M and Vickers, P. (2018) Sonification of network traffic flow for monitoring and situational awareness. PLoS ONE, 13(4): 1-31.

Dombois, F. (2001). Using audification in planetary seismology. In Proceedings of the 7th International Conference on Auditory Display (ICAD), (Helsinki,Finland, July 29th-August 1), 227-230.

Doshi-Velez, F. and Kim, B. "Towards a rigorous science of interpretable machine learning," arXiv, 2017. [Online]. Available: https://arxiv.org/abs/1702.08608

Dubus G, Bresin R (2013) A Systematic Review of Mapping Strategies for the Sonification of Physical Quantities. PLOS ONE 8(12): e82491.

Edward, C and Ville, P. (2003) Using multi-channel spatialization in sonification: A case study with meteorological data in Proceedings of the 9th International Conference on Auditory Display (ICAD), Boston, MA, July 7-9, 2003.

Eldersvelt, D (2019) Data Sonification in Power BI, Data Veld. https://dataveld.com/2019/08/03/data-sonification-in-power-bi/. Last access 19th Jan 2021.

Emmerson, S (ed.) (1986) The Language of Electroacoustic Music. London: Palgrave McMillan.

Everest Alton, F (1996) Manuale di Acustica. Concetti fondamentali - Acustica degli interni. Milano: Hoepli.

Fiebig, A, Jordan, P, Moshona, C C. (2020) Assessments of Acoustic Environments by Emotions – The Application of Emotion Theory in Soundscape. In Frontiers in Psychology, 20 Nov 2020.

Flowers, J (2005) Thirteen years of reflection on auditory graphing: Promises, pitfalls, and potential new directions in Proceedings of ICAD 05-Eleventh Meeting of the International Conference on Auditory Display, Limerick, Ireland, July 6-9, 2005.

Foo, B. (2012) https://datadrivendj.com/tracks/refugees/ Last access October 27th 2019.

Foo, B. (2015) https://datadrivendj.com/tracks/subway/ Last access January 22nd, 2020.

Franinović K and Serafin S (2013) Sonic Interaction Design. Cambridge, Massachusetts: MIT Press.

Frascara, J (2004) Communication Design: Principles, Methods, and Practice. New York: Allworth Press.

Friedman, K. (2000) Creating design knowledge: from research into practice in Proceedings of

IDATER 2000 Conference, Loughborough: Loughborough University.

Gaver, B and Bowers, J. (2012) Annotated portfolios. interactions 19, 4 (July + August 2012), 40–49.)

Gaver, B., Dunne, T., and Pacenti, E. (1999). Design: Cultural probes. Interactions 6 (1): 21-29.

Gaver, W (1993) What in the World Do We Hear?: An Ecological Approach to Auditory Event Perception, Ecological Psychology, 5:1, 1-29.

Gaver, W (1989) The SonicFinder: an interface that uses auditory icons. Hum.-Comput. Interact. 4, 1 (March 1989), 67–94.

Gaver., W (1986) Auditory icons: using sound in computer interfaces. Hum.-Comput. Interact. 2, 2 (June 1986), 167–177.

Gordo, L, Brotons, L, Herrando, S, Gargallo, G. (2020) Rapid behavioural response of urban birds to covid-19 lockdown. http://hdl.handle.net/10261/220481

Guillarme, A (2011) Intelligent Auditory Alarm. In Hermann, T., Hunt, A., Neuhoff, J. G., editors, The Sonification Handbook, Chapter19, pp. 493–508. Logos Publishing House, Berlin, Germany.

Hermann, T. (2002), Sonification for Exploratory Data Analysis. PhD Thesis, Technische Fakultät, Bielefeld University, Bielefeld.

---------------- (2008), Taxonomy and Definitions for Sonification and Auditory Display. In Proceedings of the 14th International Conference on Auditory Display. Paris, France.

Hildebrandt, T. and Rinderle-Ma, S. (2015), Server sounds and network noises. In 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), Gyor, 2015, pp. 45-50.

Hildebrandt, T. Hermann, T. and Rinderle-Ma, S. "A sonification system for process monitoring as secondary task," 2014 5th IEEE Conference on Cognitive Infocommunications (CogInfoCom), Vietri sul Mare, 2014, pp. 191-196,

Hermann T, Hildebrandt T, Langeslag P, Rinderle-ma S. (2015) "Optimizing Aesthetics and Precision in Sonification for Peripheral Process-Monitoring" in Proceedings of the 21st International Conference on Auditory Display (pp.317-318).

Hermann, T., Hunt, A. and Neuhoff, J. G., eds. (2011) The Sonification Handbook. Berlin, Germany: Logos Publishing House.

Hermann, T. and Ritter, H. (1999) Listen to your data: Model-based sonification for data analysis. In G. E. Lasker,editor, Advances in intelligent computing and multimedia systems, pp. 189–194, Baden-Baden, Germany,1999. Int. Inst. for Advanced Studies in System research and cybernetics.

Hilmann, N and Pauletto, S (2014) The Craftsman: The use of sound design to elicit emotions. The Soundtrack 7(1).

Hildebrandt, T., Hermann, T., & Rinderle-Ma, S. (2014). A sonification system for process monitoring as secondary task. In 2014 5th IEEE Conference on Cognitive Infocommunications (coginfocom) (pp. 191–196).

Hildebrandt, T Hermann, T Rinderle-Ma, S. (2016) Continuous sonification enhances adequacy of interactions in peripheral process monitoring. In International Journal of Human-Computer Studies, Volume 95, Pages 54-65.

Hogan, T. and Hornecker, E. (2016) Feel it ! See it ! Hear it ! Probing Tangible Interaction and Data Representational Modality., Des Res Soc., (Fort Lauderdale, Florida, 5-10 2003): 1-13.

Hutchinson, H., Mackay, W., Westerlund, B., et al. (2003), Technology probes: Inspiring design for and with families. In: CHI03 Proceedings of the Conference on Computer-Human Interaction, 17-24.

Iber, M., Lechner, P., Jandl, C. et al. Auditory augmented process monitoring for cyber physical production systems. Pers Ubiquit Comput (2020).

Jeon M, Walker BN, Barrass S. Introduction to the Special Issue on Sonic Information Design: Theory, Methods, and Practice, Part 2. Ergonomics in Design. 2019;27(1):4-4.

Jeon, M. (2015). An exploration of semiotics of new auditory displays: A comparative analysis with visual displays. Proc. of the 21st Int. Conference on Auditory Display (ICAD2015), 109-112.

Jeon, M and Y. Sun, “Design and evaluation of lyricons (Lyrics + Earcons) for semantic and aesthetic improvements of auditory cues,” in Proc. of Int. Conf. (ICAD2014), New York, USA, 2014.)

Jeon, M and B. N. Walker, “Spindex”: Accelerated initial speech sounds improve navigation performance in auditory menus. Proc. of Int. Conf. (HFES2009), pp. 1081-1085, 2009.)

Johnson, M. (2007). The meaning of the body. In Developmental perspectives on embodiment and consciousness (pp. 35–60). Psychology Press.

Jovchelovitch, S and Bauer, M W. (2000) Narrative Interviewing. In: Bauer, M. W., & Gaskell, G. (Eds.), Qualitative Researching with Text, Image and Sound: A Practical Handbook (pp. 57-74). London: SAGE.

Kramer, G. (1994), An introduction to auditory display. In Gregory Kramer, (ed.), Auditory Display, volume XVIII of Santa Fe Institute, Studies in the Sciences of Complexity Proceedings, (Reading, MA, Addison-Wesley), pages 1–78.

Kramer, G., Walker, B., Bonebright, T., Cook, P., Flowers, J., Miner, N., Neuhoff, J., Bargar, R., Barrass, S., Berger, J., Evreinov, G., Fitch, W., Grohn, M., Handel, S., Kaper, H., Leivkowitz, H., Lodha, S., Shinn-Cunningham, B., Simoni, M. and Tipei, S. (1999). The sonification report: Status of the field and research agenda. Report prepared for the National Science Foundation by members of the International Community for Auditory Display.

Kuckartz, U. (2014). Qualitative text analysis: A guide to methods, practice & using software. London: SAGE Publications Ltd.

Kuhn, T (1962) The structure of scientific revolutions. Chigao: University of Chicago Press. Edition 2012.

ICS-CERT (2016). “NCCIC/ICS-CERT year in review: FY 2015.” Report No. 15-50569, 558 U.S. Department of Homeland Security – Industrial Control Systems-Cyber Emergency 559 Response Team, Washington, D.C.

Lakoff, G. and Johnson, M. (1981). Metaphors we live by. Chicago: University of Chicago Press. Chicago (2003 ed.).

Latour, B. (1986) Visualisation and Cognition: Drawing Things Together in H. Kuklick (editor) Knowledge and Society Studies in the Sociology of Culture Past and Present, Jai Press vol. 6, pp. 1-40.

Latour, B. and S. Woolgar (1979) Laboratory Life : The Social Construction of Scientific Facts. London: Sage.

Lau, A. and Vande Moere, A. (2007) Towards a Model of Information Aesthetics in Information Visualization. In Proceedings of the 11th International Conference Information Visualization (IV ’07). (IEEE Computer Society, USA) 87–92.)

Lenzi S., Ciuccarelli P., Liu H., Hua Y. (2020) Data Sonification Archive. http://www.sonification.design. Last accessed: 22/02/2021.

Lenzi S, Ciuccarelli P. Intentionality and design in the data sonification of social issues. Big Data & Society. July 2020

Lenzi, S., Terenghi, G., Taormina, R., Galelli, S., Ciuccarelli, P. (2019) Disclosing cyber-attacks on water distribution systems. An experimental approach to the sonification of threats and anomalous data. In Proceedings of the 25th International Conference on Auditory Display, (June 23-27, Northumbria University). 125-132.

Leplâitre G and McGregor, I (2004) How to tackle auditory interface aesthetics? Discussion and case study in Proceedings of the 10th International Conference on Auditory Display (ICAD2004).

Manovich, L. (2008) Introduction to Info-Aesthetics. http://manovich.net/content/04-projects/060-introduction-to-info-aesthetics/57-article-2008.pdf (accessed 31st March 2020).

Masud, L., Valsecchi, F., Ciuccarelli P., Ricci, D. & Caviglia, G. (2010) “From Data to Knowledge - Visualizations as Transformation Processes within the Data-Information-Knowledge Continuum," 2010 14th International Conference Information Visualisation, London, 2010, pp. 445-449.

Mauri, M., Elli, T., Caviglia, G., Uboldi, G., & Azzi, M. (2017). RAWGraphs: A Visualisation Platform to Create Open Outputs. In Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter (p. 28:1–28:5). New York, NY, USA: ACM.

Mcfarlan, R (2016). The Secrets of the Wood Wide Web. In The New Yorker. https://www.newyorker.com/tech/annals-of-technology/the-secrets-of-the-wood-wide-web (last access 21st January 2021).

McGookin, D and Brewster, S. (2011) Chapter 14. Earcons. In Hermann, T., Hunt, A. and Neuhoff, J. G., eds. The Sonification Handbook. Berlin, Germany: Logos Publishing House.

Hazitek (2018). Memoria de Proyecto. Private document.

Murch, W (1994) Introduction in Chion, M The Audivision: Sound on screen. New York, U.S.: Columbia University Press.

Muratovski, G (2015) Research for Designers. A guide to methods and practice. London: Sage Publishing ltd.

Murphy, D., Neff, F., 2011. Spatial Sound for Computer Games and Virtual Reality. In: M. Grimshaw, ed. 2011. Game Sound Technology and Player Interaction: Concepts and Developments. Hershey PA : Information Science Reference. pp.287-312.

Nesbitt, K. V. and Barrass, S. (2002) Evaluation of a Multimodal Sonification and Visualization of Depth Market Stock Data in Proceedings of the 2002 International Conference on Auditory Display, Kyoto, Japan, July 2-5, 2002.

Nesbitt, K., & Barrass, S. (2004). Finding Trading Patterns in Stock Market Data. IEEE Computer Graphics and Applications, 24(5), 45 55.

Nees, M. (2019) Eight components of a design theory of sonification in Proceedings of 25th International Conference on Auditory Display (ICAD 2019) 23-27 June 2019, Northumbria University, Newcastle upon Tyne, UK.

Norman, D (2004) Emotional design: Why we love (or hate) everyday things. New York: Basic Books.

Offenhuber, D and Telhan, O (2016) Indexical Visualization—the Data-Less Information Display. In Ubiquitous Computing, Complexity and Culture, edited by Ekman, U, Bolter, J D, Diaz, L, Søndergaard,M and Engberg, M. 288–303. New York: Routledge.

Ostfeld, A., et al. (2012). Battle of the water calibration networks. In J. Water Resour. Plann. Manage., 10.1061/(ASCE)WR.1943-5452 .0000191, 523–532.

Oswald, D (2012). Non-Speech Audio-Semiotics: A Review and Revision of Auditory Icon and Earcon Theory in Proceedings of the 18th International Conference on Auditory Display, Atlanta, GA, USA, 18-21 June 2012.

Pauletto, S (2014), Editorial in The New Soundtrack. Special Issue Perspectives on Sound Design 4(2), pp. v–vi. Edinburgh University Press.

Peirce, C S (1902) Logic as semiotic: The theory of signs. Philosophical Writings. (2011 ed). Mineola, NY: Dover Publications.

Petocz, A, Keller, P and Stevens, C. (2008). Auditory warnings, signal-referent relations, and natural indicators: Rethinking Theory and application. In Journal of Experimental Psychology: Applied , vol. 14, no. 2, Jun 2008, 165-178.

Roddy, S. & Bridges, B. (2016). Sounding Human with Data: The Role of Embodied Conceptual Metaphors and Aesthetics in Representing and Exploring Data Sets. In Proceedings of MusTWork16 – Music Technology Workshop 2016. (10-16 June, Dublin, Ireland).

Roddy, S. (2015) Embodied Sonification. PhD Thesis, University of Dublin, Trinity College.

Roddy, S and Furlong, D (2015) Sonification listening: An empirical embodied approach in Proceedings of the 21st International Conference on Auditory Display (ICAD 2015). July 6-10, 2015, Graz, Styria, Austria.

Roddy, S (2018) Ambient Data Monitoring w/Generative Music Systems using EC & ML Techniques. In Proceedings of Computer Simulation of Musical Creativity(CSMS2018)., 20th-22nd August, University College Dublin.

Roddy, S. & Furlong, D. (2014). Embodied Aesthetics in Auditory Display. Organised Sound. 19(Special Issue 1. Sonification):7-77.

Roetzel, P.G. (2019), Information Overload in the Information Age: A

Review of the Literature from Business Administration, Business Psychology, and Related Disciplines with a Bibliometric Approach and Framework Development. Bus Res 12, 479-522.

Rönnberg, N. (2017). Sonification Enhances Perception of Color Intensity. In Proceedings of IEEE VIS 2017At: Phoenix, AZ, USA, 2017.

Roth R.E. (2017). Visual Variables. In International Encyclopedia of Geography: People, the Earth, Environment and Technology (eds D. Richardson, N. Castree, M.F. Goodchild, A. Kobayashi, W. Liu and R.A. Marston).

Shapin S and Schaffer, S. (1985) Leviathan and the Air-Pump : Hobbes, Boyle, and the Experimental Life. Pinceton, New Jersey: Princeton University Press. Edition: 1989.

Schaeffer P. (1966) Traité des objets musicaux, pp. 91-99. Paris: Seuil.

Schaefer M R (1977) The tuning of the world. New York : A.A. Knopf, 1977.

Ricci D, Ciminieri, D, Mauri, M, Ooghe-Tabanou, B, Plique, G, Verjat, B (2015) Seealsology. Link: https://medialab.sciencespo.fr/en/tools/seealsology/ Last accessed 28th January 2021.

Sendelbach S and Funk M. (2013) Alarm fatigue: a patient safety concern. AACN Adv Crit Care. 2013 Oct-Dec;24(4):378-86; quiz 387-8.

Snow, C P. (1959) The two cultures. Cambridge: Cambridge University Press; Reissue edition (March 26, 2012).

Supper, A. (2012) Lobbying for the ear: the public fascination with and academic legitimacy of the sonification of scientific data. PhD Thesis, Univ. Pers Maastricht.

Susini,P, Houix, O, Misdariis, N. (2014) Sound design: an applied, experimental framework to study the perception of everyday sounds. The New Soundtrack, Edinburgh University Press, 4 (2), pp.103-121.

Scaletti, C. (1994) Sound Synthesis Algorithms for Auditory Data Representations. In Auditory Display. Sonification, Audification, and Auditory Interfaces, edited by G. Kramer. Reading: Addison-Wesley Publishing Company, 223-251.

-------------- (2018) Sonification ≠ music in Alex McLean and Roger Dean (eds.), The Oxford Handbook of Algorithmic Music. New York: Oxford University Press, 2018.

Scaletti, C. and Craig, A. B. (1991) Using sound to extract meaning from complex data. in Proc. SPIE 1459, Extracting Meaning from Complex Data: Processing, Display, Interaction II, (1 June).

Schön, D A (1992) The Reflective Practitioner : How Professionals Think in Action. London: Taylor & Francis. Second Edition: 1995.

Sobliye, G. and Mortada, L. (2017). Uneasy listening: El subutilizado arte de sonificar datos para contar historias de injusticias sociales. Revista Diseña. 11. 202- 219.

Sonnenschein, D (2001) Sound Design: The Expressive Power of Music, Voice and Sound Effects in Cinema. San Francisco: Michael Wiese Production.

Strauss, A. and Corbin, J. (1998). Basics of qualitative research: Techniques and procedures for developing grounded theory (2nd ed.). London: Sage Publications, Inc.

Taormina, R., Galelli, S., Tippenhauer, N. O., Salomons, E., & Ostfeld, A., “Characterizing cyber-physical attacks on water distribution systems in Journal of Water Resources Planning and Management, 143(5), 04017009, 2017.

Taormina, R., Galelli, S., Tippenhauer, N. O., Salomons, E., Ostfeld, A., Eliades, D. G., ... Ohar, Z. “Battle of the Attack Detection Algorithms: Disclosing cyber-attacks on water distribution networks.” in Journal of Water Resources Planning and Management, 144(8), 2018.

Terenghi, G (2019) Sonificazione di dati e sicurezza informatica nelle reti di distribuzione idrica. Il caso C-Town. Master’s degree thesis. Politecnico di Milano, April 2019. http://hdl.handle.net/10589/147171

The Design Exchange (2020), http://www.thedesignexchange.org. Last accessed 25th January 2021.

Truax, B (1978) The Handbook of Acoustic Ecology. A.R.C. Publications : A.R.C. The Aesthetic Research Center : World Soundscape Project ; Vancouver ; Burnaby, B.C.

Vande Moere, A. and Purchase, H.C. (2012) On the role of design in information visualization. Information Visualization, 10(4), pp. 356-371.

Vickers, P. (2011). Sonification for process monitoring. In Hermann, T., Hunt, A., Neuhoff, J. G., editors, The Sonification Handbook, Chapter18, pp. 455–491. Logos Publishing House, Berlin, Germany.

Vickers, P. and Hogg, B. (2006). Sonification Abstraite/Sonification Concr\ete: An 'Aesthetic Perspective Space' for Classifying Auditory Displays in the Ars Musica Domain in Proceedings of the 12th International Conference on Auditory Display (ICAD2006). June 20-23, London, UK.

Vickers, P. and Barrass, S. (2011) Sonification design and aesthetics. In Hermann, T., Hunt, A., Neuhoff, J. G., editors, The Sonification Handbook, Chapter 7, pp. 145–171. Logos Publishing House, Berlin, Germany.

Vickers, P., Laing, C., Debashi, M., & Fairfax, T. (2014). Sonification Aesthetics and Listening for Network Situational Awareness. In Proceedings of SoniHED-Conference on Sonification of Health and Environmental Data, (York, U.K., 12th September).

Vickers, P and Hogg, B. (2006). Sonification abstraite/sonification concrete: An 'aesthetic persepctive space' for classifying auditory displays in the ars musica domain in Proceedings of the 12th International Conference on Auditory Display (ICAD), London, UK, June 20-23, 2006.

Vogt K (2011) A quantitative evaluation approach to sonifications in Proceedings of the 17th International Conference on Auditory Display (ICAD 2011).Budapest, Hungary.).

Vogt, K. and Höldrich, R., (2010). Metaphoric Sonification Method-Towards the Acoustic Standard Model of Particle Physics. In Proceedings of the 16th International Conference on Auditory Display (ICAD2010). (June 9-15, Washington D.C., U.S.A.), 271-278.

Walker, B and Nees, M (2011) Theory of Sonification in Hermann, T., Hunt, A. and Neuhoff, J. G., eds. The Sonification Handbook, Chapter 2. Berlin, Germany: Logos Publishing House.

Walker, B N Nance, A and Lindsay, J. (2006). Spearcons: Speech-based earcons improve navigation performance in auditory menus in Proc. of the International Conference on Auditory Display (ICAD2006), 19-23 June, London Queen Mary University.

Walker, B., and Kramer, G. (2004). Ecological Psychoacoustics and Auditory Display. In Ecological psychoacoustics, edited by J. G. Neuhoff. San Diego: Elsevier.

Walker, B. N. and Kramer, G. (2005). Sonification Design and Metaphors: Comments on Walker and Kramer, ICAD 1996. In ACM Trans. Appl. Percept. 2(4): 413–417.

Walker, B and Kramer, G (1996) Mappings and Metaphors in Auditory Displays: An Experimental Assessment. ACM Transactions on Applied Perception, 2(4), 407–412.

Walker, B and Lane, D (2001) Psychophysical scaling of sonification mappings: A comparison of visually impaired and sighted listeners in Proceedings of the 7th International Conference on Auditory Display. July 29th - August 1st, 2001. Helsinki University of Technology.

Walker, B. and Cothran, J. T. (2003) Sonification sandbox: A graphical toolkit for auditory graphs In Proceedings of the 9th International Conference on Auditory Display (ICAD), Boston, MA, July 7-9, 2003

Weizman, E. (2017) Forensic Architecture: Violence at the Threshold of Detectability, Princeton, NJ: Zone Books, Princeton University Press.

Wikipedia The free encyclopedia (2021) Sonic Interaction Design. https://en.wikipedia.org/wiki/Sonic_interaction_design Last accessed 28th January 2021.

Wirfs-Block, J., Fam, A., Devendorf, L. and Keegan, B. (2021) Examining Narrative Sonification: Using First-Person Methods to Translate Radio Production to Interaction Design in Transactions on Computer-Human Interaction – Special Issue on First Person Research in HCI, In press.

Worral, D. (2009) Sonification and Information: Concepts, instruments and techniques. PhD Thesis, University of Canberra.

Worral, D., Blystar, M., Barrass, S. and Dean, R. (2007) Sonipy: The Design of an Extendable Software Framework for Sonification Research and Auditory Display in Proceedings of the 13th International Conference on Auditory Display, Montréal, Canada, June 26-29, 2007.