‘MUSIC OF THE PEOPLE’: MUSIC FROM DATA AS SOCIAL COMMENTARY

Rob King
Music Dept., Durham University,
Palace Green,
Durham, DH1 3EP, United Kingdom
robert.g.king@durham.ac.uk

ABSTRACT

Data-music reflects the ubiquity of data in modern society. Composers have not engaged widely with the opportunities opened up by this, despite the chance to overcome a gulf between academic art music and social engagement. Their reluctance might be traced to the challenge of reconciling abstract data and concrete sound, in political implications, and in technological barriers in computer music. The present paper argues that socially relevant music composition for the 21st century can adopt a programme of sonification grounded in politically acute data. As examples of such practice, two compositions are discussed founded upon US and UK social data sets, and realised via the SuperCollider programming language. The consequences for the composer of new music are further discussed from political and musicological angles, with the ‘purpose’ of writing such music analysed from the perspective of various commentators.

1. INTRODUCTION

There are more data collected on humans today than at any period in history. These data are given freely, by people answering questionnaires and ‘liking’ photographs, but are also harvested by corporations logging sites and products we look at online, or by government-sponsored security firms watching us through the lenses of countless cameras. It is the responsibility of global citizens to be aware of the data collected on them, and yet too often we are keen to ‘accept all terms and conditions’ without proper scrutiny [1]. Presenting data as art can provide a fresh opportunity to examine this datopia, with the potential to plant questioning seeds in the minds of those whom the data concern. Composers have an opportunity to engage with and critique contemporary society; in an internet-based world where clicks are currency, an artistic representation of society’s greatest social and economic force can offer food for thought for the demographics from whom the data are collected [2].

Moreover, there exists a disconnect between society and classical music – concert attendance is decreasing, record sales low, and contemporary art music perceived to be deliberately difficult and arcane [3, 4]. By no means does this paper suggest a ‘solution’ to this ‘problem’, but it does propose a method by which contemporary art music can reconnect with a wider populace.

There have long been efforts to integrate data into composition work. These stem from the efforts of composers to treat particular data as compositional material, such as Iannis Xenakis and stochastic functions, Charles Dodge and geomagnetic data, or Natasha Barrett and weather systems, often with the aid of computers (in Xenakis’ case, his first access to computers was in 1962). Xenakis viewed the computer as a means by which he could write mathematically driven music - he was not, as I am, interested primarily in the computer as a sociological phenomenon, and its use as a signifier of the information age in the creation of art [5]. Social, political, environmental, and personal data are often rich and interesting sources, to be sure, but carry with them much extramusical baggage. Xenakis’ music was not based on societal phenomena; his sonifications were mathematical, platonic conceptions. Indeed, there exist comparatively few composers writing about political issues by using data from humanity.

One composer who does treat sociological data is R. Luke DuBois. He views his work as primarily political, as opposed to technological, “using the tools of a particular established medium to critique it” [6]. It is this politicised approach to writing music using data that this paper seeks to consolidate. Other musicians have further refined the field to sonifying social data; for instance, De Campo and Egger De Campo in 1998, using data on executions in the USA since 1977 [7].

The case studies examined in this paper do not present data in a manner in which the data are legible as information, pace the model of Vickers and Hogg [8]. The music presented has had certain compositional processes applied to it, so that the data forms a basis for the music; the translation of data-parameters to musical ones has been undertaken, primarily, as a musical endeavour, with musical considerations. As such, sonifications of this sort have a slight tendency towards ars informatica, but the decisions of the composer mean that they retain the critical element that classifies them as musical - that they are organised by a composer, first and foremost. This is a valid measure; art is a response to stimuli, but those responses are filtered through the personality of the artist. Thus, it is reasonable for someone who is creating art to manipulate that art in line with their views on the world; to do so in the field of data science, or to claim artistic representations of data are scientific would be unjustified.

Presented in this paper are a number of works for a combination of computer and live performers. Incorporated are considerations about the sources of, practical employment of, and political implications of the use of data in composition. The works are complemented by SuperCollider code examples, and are further contextualised by reference to Vickers & Hogg’s model, and yet further to the wider political context in which they have been created [8].
2. COMPOSITIONAL EXAMPLES

There exist, as has been discussed, more data than ever on the subject of humanity and its habits. For the purposes of this paper, data has been selected from verifiable sources, such as the Federal Bureau of Investigation in the United States of America, or the Office for National Statistics in the United Kingdom (for the first and second case studies respectively). An examination of potential future projects using data from unverifiable sources is given in Future Directions (section 3), below. Sound examples of both case studies are available at: soundcloud.com/robking-2.

2.1. Case Study #1

In the first piece, data from gun license background checks carried out in the US between November 1998 and January 2019 were sonified. This sonification involved mapping state by state monthly gun license checks (represented in fig.1 on the y-axis) to pitch.

The passage of time, represented in the graph by the x-axis, is mimicked by the composition (in which, instead of years, the data is presented in a matter of minutes). The choice of presentation timescale is the most critical artistic decision to be made, alongside whether to present the data with minimal compositional interference, hoping that the audience receive a performance as close to ars informativa as possible, or whether to react to the data emotionally, and imbue the character of the resulting piece with this reaction.

In this case, the decision was made to have the perceived tempo of the piece increase over time. This complemented the data - generally, there was an upward trend in gun sales over time, so this was matched by the increase in perceived tempo. It also made the piece sound much more complex as it continued - this, too, was a deliberate choice, as the generally recognisable cyclicity of the first 10 years of data becomes disrupted and disfigured, especially around the time of a large upwards spike in 2013.

The piece features a clear accelerando effect, as the duration of the notes decreases following a quadratic function. The reason for such an accelerando is the composers’ perception of the quickening frequency of national disasters in the USA brought about by gun violence around the Autumn and Winter of 2017.

It is not only more compelling to highlight the increased frequency of gun license background checks; it is in fact necessary. This is in accord with Cardew’s assertion that “[the composer] must demand works that relate directly to the issues and struggles and preoccupations of the present […]” [9]. As such, the piece draws not only on the data, as published by the FBI, but the reporting in the media that mass shootings in the USA are becoming more frequent [10], and attempts to highlight this in a sonic fashion.

By outputting a MIDI File from SuperCollider based on this reading of the data, and introducing this MIDI to Sibelius, a score was made. This score was designed to be played by a pianist, but, due to its complexity, this is a difficult task. This is an anticipated side-effect, and is more informed by compositional theories concerning New Complexity as a genre of contemporary music, than it is by computer science or data analytics, and is intended to make the experience of performing the work stressful for the performer. Simply put, the intention is to depict a difficult, contentious subject in a difficult manner with a difficult aesthetic, in line with the composer’s views on said subject, so as to align the work as ars musica above ars informativa.

Lastly, to reference Vickers & Hogg’s other axis (concrete-abstract), it should be noted that, by translating the MIDI data to a score designed to be performed by solo pianist, the work is immediately driven towards the abstract end of the scale, where the sounds generated during composition do not represent the sounds of the data. For instance, a less abstract composition, using the same gun licensing data, could use general MIDI sound 128 - the sound of a gunshot. Alternatively, gun license background checks could be sonified with the sound of a cash register (as the checks are required whenever a customer wishes to purchase a firearm), panned left or right depending on longitude data from the state in which the check was carried out.

Analysis of compositional method of this type gives composers the opportunity to write music, in the knowledge of where their work ranks in terms of abstraction, and therefore gives them tools with which they might self-assess their composition and its nature. The utilisation of a concrete-abstrait axis also encourages composers to think about the sounds they use in sonifications, thereby giving them an opportunity to address the allegation (addressed in the introduction) that their music is wilfully obscure.

2.2. Case Study #2

This piece is formed from self-report data published on loneliness in young people in Britain, aged between 16 and 24 [12]. These data are presented in a more detailed manner than those in the previous example, in that they are not presented as a measure of one group over time; there are three groups (those who often, sometimes/occasionally, or rarely/never feel lonely, respectively) which are presented as percentages of the population. Instead of presenting these groups over time, the data are from a single survey in May and June 2018. They are further categorised in the downloadable file by gender, race, and other relevant markers.

1 Although, this complication makes the connection between those from whom data is harvested and the music that derives from these data more difficult to establish. This is discussed further in section 3.
This means that, unlike Case Study #1, there are no audible trends over time. Similarly, there cannot be a presentation of the data using ‘soundalike’ MIDI, as there is no real-world sound associated with the data. Indeed, it could be said that loneliness is associated with an absence of sound. This is where considerations as a composer must come to the fore - how can a piece based on loneliness data achieve a clear aesthetic, when there is no way that the audience could notice trends in the data, due to the way that they have been collected and collated?

This ambiguity is intentional, and aimed at potential criticism of music generated by sonifying data. Specifically, it addresses the comprehensibility and audibility of data, and presents the notion that the individual data need not be understood for the listener to understand the overall aesthetic of the piece as music [13]. Whilst the data may not be legible, at a datum-by-datum level, the overall affect of the piece can still be communicated by other factors involved in the presentation of the material.

This is the reason for the final format of the piece: the simultaneous presentation of music performed by laptop and by solo pianist. The solo pianist is (necessarily) alone on the stage - a decision taken, not because of its implication in the data, but because of the compositional desire to communicate loneliness. This is yet another reason why, when presented ‘in a vacuum’ sonified data can appear obscure, and its meaning oblique, but when taken holistically, a concert setting can communicate the crux of the work. Other decisions have been taken to communicate the lonely aesthetic still further - decisions that the data could not possibly have dictated. For instance, the pianist is directed to face away from the audience.

Making choices to communicate the essence of the piece is also the reason for compositional choices made in the musical material. Where no meaningful impression of the data is left by its translation to pitch, it is the job of the composer to communicate ‘beneficial’ meaning to the audience, in order that it might ‘raise their consciousness’ [9].

The choice of an increasing tempo, as per the first case study, was eschewed in favour of a more ‘lonely’ aesthetic, with irregular beats, making a tempo difficult to determine for the listener. Thus, the piano solo score is sparse, the feeling of being without a tempo contributing, once more, to the effect of isolation. See fig. 2 for an example from the score given to the pianist - note the long, sustained, comparatively infrequent notes.

By the same token as the decisions surrounding tempo, it would not have made sense to present the material as one long, continuous line (as in the first case study) as it is not organised chronologically. There are three columns of data presented, recorded simultaneously from three self-identifying groups (those who never, sometimes, or often feel lonely). Thus, these are sonified simultaneously, so as to better reflect the manner in which the data were presented originally.

The choice of instruments has been discussed already, but the individual design of these instruments is also a key consideration. There are two synthisising instruments, created in SuperCollider, which have been designed so as to mimic the timbre of a piano with a few key differences. Whilst the lonely pianist plays the data derived from those who often or always feel lonely, the first synthesised piano performs that which is derived from those who sometimes or occasionally feel lonely. The SuperCollider plugin ‘MdaPiano’ was used, and its parameters manipulated to be comparatively close to that of a real piano (the parameters of decay and reverb, for instance, could be modified in a performance to more closely mimic the live piano). The second sound sonifies the data from those who identify as rarely or never feeling lonely. The SuperCollider plugin ‘MdaPiano’ was used, and its parameters manipulated to be comparatively close to that of a real piano (the parameters of decay and reverb, for instance, could be modified in a performance to more closely mimic the live piano). The second sound sonifies the data from those who identify as rarely or never feeling lonely. Thus, it is designed to sound somewhat different from a piano. It retains, for example, a quick attack, but its decay and resonance values make it sound dissimilar to a piano over the course of the envelope of a note. It is also different from MdaPiano in that it is formed from a combination of harmonics.

These three instruments are tiered in such a hierarchy as a play on the already-established bias towards a visual presentation of data over the sonic. The piece acts as a kind of thought-experiment for the listener. The synthesising pianos present similar material to the lone pianist (again, so that no pulse can be easily identified, nor any tonal centre, nor thematic material), thereby leading the audience to the main difference between the sounds - their timbre.

The hierarchy is conceived so that the most piano-like sound (the real piano) is given the most attention - both in terms of staging and in terms of the complexity and ‘realism’ of its envelope. The sound dedicated to the data from those who only sometimes feel lonely is the next-most similar to a piano, as a metaphor for the implied empathy with the pianist. By the same token, the second (most unrealistic) synthesised piano is symbolic of the concept of it having the least in common with the live pianist. Its sounds are noticeably dissimilar to the real piano, metaphoric of the idea that the audience should empathise with the live pianist the most, out of the three sounds.

The ‘disembodiment’ of the two synthesised piano sounds intentionally shifts focus away from them. The simultaneous presentation of sonic and visual stimuli is further discussed later, but the removal of the visual cue of a performer from the synthesised piano sounds in this case reframes the solo pianist as the primary source of engagement in the performance.

The overall effect is designed to encourage the audience to associate with the pianist more than the synthesised sounds, in order to present a figurative ‘cure’ for the loneliness the pianist feels. This is achieved by the selection of timbres and staging cues; the data, in their raw form, could not communicate this, so it is incumbent on the composer to do so instead. The composer devises a musical response to data, not a scientific sonification, and there is therefore a degree of justifiable ‘artistic license’.

![Fig. 2: An excerpt from the score generated for solo pianist; note the long durations and sparse texture.](image-url)
3. WIDER PERSPECTIVES

The use of data in modern society is universal; in science, marketing, politics, and industry, its use is critical. Contemporary art, however, does not reflect this ubiquity. Political music has existed for countless years, but given the comparatively recent emergence of data as a force in the digital age, a composer must now reconcile themselves with the consequences of using data that is, at its core, derived from humanity. Musicologists have critically interrogated the composition of music using techniques derived from traditional Western genres, and indeed the place of abstract music as an art form has been problematised, by commentators such as Cornelius Cardew and Susan McClary.

There are a panoply of different projects that data use of this sort could produce. One of particular interest is that of the installation. Installations offer a number of exciting opportunities for composers and artists, in that they can transcend the musical/informatica axes of Vickers & Hogg, by presenting abstract sound and highly representative data simultaneously, as in work of Ryoji Ikeda (for instance, his audiovisual work datamatics [ver. 2.0] [14].

Furthermore, there still exists a bias toward trusting visually displayed data more than sonified [15]. Simultaneous presentation of visual and sonic data provides an opportunity to the viewer to strengthen the bond between what they see and what they hear [16]. Ikeda himself touches on this, commenting on datamatics [ver. 2.0]: “I like the invisible phenomena in sound. Data you can see as a result on the display monitor, but the concept of data is so abstract you can’t touch it.” [17].

As a composer, though, I am concerned with presenting data in an artistic and immersive manner. It is known that, when one of the senses is neglected, “information and meanings derived, and the affective engagement invoked, will be decreased; everything from realism to user satisfaction, from dimensionality to ease of use, will suffer unacceptably.” [15]. Installations, by offering a parallel representation of sonified and visualised data, can combat pro-graphic bias in the public’s eye. They can provide a more stimulating artistic experience, where the communication of data and the audience’s emotional response are both improved. This can lend installations an element of the sublime, where the constituent parts of the work meld together for an overall effect, ‘greater than the sum of its parts’ [18].

If the aim is art, there are those who have employed algorithms to aid creation and tailored the outcomes according to their needs [19, 20]. A purely formalist approach is promoted by those concerned with the scientific sonification of data, but is not always adopted by artists - certainly not in the instances of the case studies above - and this raises ethical questions. Is artistic interpretation of data, or the alteration of a sonic product of said data, essentially a lie? Provided a disclaimer that the product is artistic as opposed to scientific or factual is given, there is no ethical dilemma, but there are always considerations like these to be taken into account when using (potentially contentious) data from human subjects.

However, if the aim is to report on data to the wider public, via the vehicle of the media, for instance, there are undercurrents and tensions. The idea of collecting verifiable data in the case studies has been discussed, but what happens when those data are not available, or are corrupted? ‘Fake news’, and its relationship with music and musicology, are topics for a different paper, but it can still be discussed how the presentation of fake news might be undertaken by a composer.

As previously mentioned, data that are sonified are inherently less trusted than those that are visualised. How could this distrust be brought to bear in a composition? The approach could be somewhat similar to the example of the second case study - by using unfamiliar or ‘mock’ instruments, that are somehow alien to the listener. However, whilst most listeners might not be familiar with the nuances of every possible synthesiser configuration, many people could be reasonably expected to have come into contact with abstract synthesised sounds due to their widespread presence in popular music. For the production of a piece based on fake news, a more radical direction is proposed: rather than a new instrument that may not be noticed, it is suggested that fake news is sonified alongside verifiable data by using two opposed tuning systems. For instance, due to its almost universal application in Western music since the mid 19th century, verified data could be sonified using just intonation, whereas other, uncommon or unique tuning systems could be sought for the fake news. This would create a parallel between the sonic dissonance created, and the cognitive dissonance we experience when we hear a ‘fact’ based on false evidence.

Wider than these positions, though, are the reasons why a composer might find themselves drawn to writing data-music. Susan McClary has argued that composers have been very frequently inspired by abstract concepts of ‘talent’ and ‘genius’, even more so than the other art forms [23]. Contrary to this position, though, is that which this paper takes: that music is “essentially a human, socially grounded, socially alterable construct” [23].

A literal interpretation is acceptable: that one’s compositions are rooted in the societal history and mutual consciousness of humanity. However, on a more practical level, the idea that music is deliberately mystified by producer and consumer alike, is the stance against which this paper stands. Reification of music above the real-life circumstances in which it is created is, in the mind of Cardew at least, something to be fought against, as it does not “relate directly to the issues and struggles and preoccupations of the present, and lead the way forward to a better society” [9].

4. CRITICISM

Individual drawbacks have been highlighted in the case studies concerning legibility of data, and the use of abstract, musical sounds in sonifications. There is wider debate about the field of sonification as it currently stands, and its drawbacks. For example, it has been argued that, in spite of efforts like ours to create socially responsible pieces of music, sonification is an expensive undertaking [18]. That being said, I contend that, actually, the key components – a computer and an internet connection, and some software with which to sonify data – are no longer prohibitively expensive. SuperCollider is free, and 9/10 UK households have internet access [23]. The issue is slightly more complex than this, however, as the software required for sonification requires a certain amount of time and learning to understand - privileges which are not afforded to everybody. There is more work to be done.  

2 For instance, there has been a recent spike in ‘fake’ albums being “released” by popstars, later to be discovered to be fan-made hoaxes [22].
do on interactive software, to “encourage casual exploration of sonification”, so as to improve access to the field [24].

More pressing than this, the issue of how to interpret data that have been sonified comes to the fore. These concerns mainly break down into two areas: cognitive or perceptual abilities of the listener, and musical or aesthetic considerations taken by the sonifiers.

By mapping data to the range of a piano proportionally, the second case study addresses one of the issues of perception: that, should this decision not have been taken, the resulting frequencies will be outside of a human’s hearing range [15]. Similarly, being able to discriminate between sounds as individual events is important if those sounds represent discrete data points. The first case study accelerates to such an extent that, whilst data are audible as individual points at the beginning, the overall effect towards the end of the piece is more aimed towards presenting a trend than an array of individual data. Moreover, there exist, and may never exist, no agreed methods for sonification in the wider community, such that a level of interpretation or familiarity with the material or method are required in order to get the most detail from a sonification [24, 15].

Musical considerations, too have been highlighted as areas of contention. Vickers and Hogg have claimed that sonifications and music are the same thing, depending on perception [8], and have highlighted how attention to musical factors (realistic timbre or familiar tonality, for example) can aid in the communication of data as sound. In both case studies above, musical factors have been brought to bear (such as the dramatic accelerando effect in the first, or the sparse texture in the second), but these are compositional choices, not derived from information inherent in the data. I contend, though, that the decision to present the pianist as alone on the stage (in the second case study), facing away from the audience, are artistic choices that, whilst not demanded by the data, are considerations that are in line with the nature of the data.

The two case studies herein are evidently not purely scientific sonification - they do not communicate quantitative, discrete data to an audience. However, through the implication of musical and extramusical compositional decisions that have been taken, a degree of the original meaning of the data is still possible to pick up, from the perspective of the listener. Where, then, does this leave the case studies in the wider context of data-music? The use of the term ‘data-music’ is not by accident: it is contended that, for cases where music is derived from data (rather than ‘pure’ sonification), ‘data-music’ is a better term, as it accounts for the music of composers who have “turned to sonification as a seed or driver for their work” [13]. By doing so, it allows for the scientific pursuit of the representation of data in sound, whilst establishing a parallel space for composers looking to the field for inspiration in their choices as artists.

5. CONCLUSION

This paper has explored the engagement of contemporary composers with data in their music, and how commentators have reconciled various usages with the abstraction of musical art. This paper has gone beyond the sonification of data, and its categorisation on an ‘information-music’ continuum, and has discussed the use of installations and other ways in which composers can use this continuum to generate ideas for new compositions, or alter current compositions. The limitations of writing music based on data are also acknowledged, in terms of communicating the underlying nature of those data, or decoding the ‘meaning’ behind the data.

Moreover, the political reasons for writing music based on data (informed by the arguments of composers and musicologists) have also been discussed. Music based on theories of tonal and post-tonal composition have their place in the canon, yet, as has been expressed, data-music is still an emergent art - this paper has sought to outline ways in which data can be used to write music, and, furthermore, to compose whilst cognisant of contemporary issues in musicology, and in wider political theatres.

6. ACKNOWLEDGMENT

With thanks to Nick Collins and Eric Egan for supervision in matters musical and political. I am also indebted to Erin Johnson-Williams, for her assistance in developing my musicological thought.

7. REFERENCES


