COMP0161 Individual Coursework Assignment 2025/26

Summary

Your task is to **sonify some data** – that is, produce an audio representation of the data that captures (some of) its salient features in an intuitive or revealing way.

Some candidate data sources are given below, but you are welcome to choose something else – indeed, we very much encourage you to select a dataset that aligns with your own interests.

Tools & Technologies

Ifat covered different approaches to sonification in her lecture of week 3, and we looked at a very simple example in that week's tutorial. You may draw on that example, but you must do your own analysis and design – do not simply reproduce the (often arbitrary) choices from the tutorial.

You are required to submit a single fixed sound file, so you will probably find **parameter-based** sonification approaches more suited to the task. If you prefer to produce a **model-based** interactive sonification you may do so, but you will need to record a "walkthrough" that you can submit.

You may reuse audio code from the tutorials, or write your own, or use other libraries. You may make use of field recordings or found sounds or existing samples. If you want to build your whole sonification out of the Wilhelm Scream or Amen Break, you can. You may use most other tools for audio or music creation, **provided the work you do with them is your own**. (See the section on GenAI for more on this.) Some potential options are mentioned in the Resources section below, but you are not required to use any of them. If you are already familiar with a particular audio application or environment, feel free to use that.

In all cases, you must **cite your sources** in your report and make clear how you have used them.

Data

You may use any data source that appeals to you, but you should look for something with sufficient variety to produce interesting results. Some possible candidate datasets from Kaggle are noted below, but you are welcome to ignore these and use something completely different. Cite and, where possible, include a link to your source(s) in your report.

- New York City Air Quality
- Global Seismic Events
- Monthly Electricity Production

Resources

A good starting point for figuring out your design is the Data Sonification Canvas by Sara Lenzi & Paolo Ciuccarelli (see also the paper about it). Their Data Sonification Archive includes many interesting sonification examples.

The Sonification Handbook is the "bible" of sonification, covering many aspects of the field. It is freely available in PDF form from that site.

There are **many** computational tools available for creating and manipulating sounds. A few are even dedicated to sonification:

- Highcharts Sonification Studio
- Data Sonifyer
- STRAUSS

(It is okay to use these, but you may find them limiting – they impose their own assumptions about what's being done, which can be quite simplistic.)

You will probably need an audio editor. If you don't have anything that you're already familiar with, Audacity is free, capable and very widely used. You probably *won't* need a "Desktop Audio Workstation" or DAW, and if you do you probably know it already. But GarageBand is free on Mac & iOS, Cakewalk is free (up to a point) on Windows & Mac, and BandLab is free online.

Most coding environments will enable you to produce audio, perhaps with the installation of additional libraries. As we've seen in the tutorials, low level generation is reasonably straightforward in Python, and libraries such as Pedalboard implement useful signal processing functionality. Web Audio is built into modern browsers and

accessible via JavaScript. Tone.js uses it to provide a handy framework for audio synthesis and control. Alternatively, p5.sound exposes Web Audio functionality for p5.js. If you're happier with C++ then JUCE is a very popular cross-platform library (commercial, but free for non-commercial and educational uses). That may be a bit heavyweight for this coursework, though.

There are also plenty of dedicated audio & music coding environments, including:

- Pure Data (Pd)
- SuperCollider
- Csound
- Cmajor
- ChucK

TidalCycles is a nice livecoding environment that uses SuperCollider as its synth engine. But if you just want to give it a try you're probably better off with Strudel, a JavaScript version that runs in your browser.

Freesound is a good source of audio samples. As always, remember to credit your sources.

Use of Generative AI

This coursework falls broadly into UCL Category 2 with respect to use of GenAI: GenAI tools can be used in an *assistive* role. However, due to the nature of the coursework, we take an extremely narrow view of what constitutes legitimate assistance. Please carefully read the notes below, and if you are even slightly uncertain about the acceptability of any intended use of GenAI in your work, **ask for clarification**.

Guiding principle: both the sonification and your report explaining it must be strictly your own work.

As we will discuss in the final lectures of the module, questions of authorship in the context of algorithmic composition and generative media can be quite fraught. The author's influence over the end product might be deliberately indirect. It is legitimate for that to be the case for your sonification – the content will necessarily be determined by the data to a significant extent.

Nevertheless, you must be able to demonstrate unambiguously that you **are** in fact the author. You must justify your choices and actively implement them. **You must not delegate any of those choices to an "AI" agent.**

Specifically:

- You **may not** use general purpose generative AI systems like ChatGPT to produce any part of your submission, either the audio or any of the report content.
- This includes all but the most trivial of textual "improvements" to your report. If the GenAI is changing the style or substance of your sentences then you are not the author. Do not ask it for suggestions and do not accept any it volunteers. (Dedicated spelling and grammar checkers are fine.)
- You **may not** use audio-specific generative AI systems like Suno or Udio that produce sound in response to a prompt or description.
- You **may** use dedicated machine learning models such as WaveNet or Magenta DDSP, **provided** that you are individually responsible for defining and implementing exactly what you do with them. You will need to document this in your report.
- You **may** use an AI coding assistant such as Copilot or Cursor to help with programming tasks, **provided** that you are individually responsible for designing and specifying the tasks being implemented. You will need to document this in your report.
- You **may** use music tools such as the "virtual drummer" in GarageBand to produce simple background accompaniment or other elements included in your audio, **provided** their role in the final product is decorative or supplementary and does not substantially contribute to the semantics of your sonification.

IMPORTANT: You must provide a statement in your report explicitly documenting all uses of GenAI, specifying the systems and/or models used and what they were used for. This statement counts towards your page and word limits. It is not worth any marks in itself.

Submission Format

Your submission must consist of **exactly 2 files**:

- An **audio file** of 60–90 seconds duration, containing the actual sonification.
  - This can be a representative clip or excerpt from a larger output, but if so the excerpt should be self-contained. For the purposes of this coursework it will be judged in isolation.
  - The audio must be submitted in a standard pulse code modulation file format, such as WAV, M4A/AAC, MP3, FLAC, AIFF. Do not use anything obscure – it is your responsibility to ensure that this will be playable by the people marking your work.
  - If your sonification process generates its output in some other format, such as MIDI, you must render it to audio before submitting.

- A **short report** in PDF format documenting your sonification. This should be no more than 600 words *and* no more 2 sides of A4.

Length limits will be strictly enforced. Audio files and/or reports that exceed the stipulated length will be cut off at the limit and marked as if the remaining material had not been submitted. There is no penalty attached to the overrun itself, but your truncated coursework will likely be incomplete and lose marks as a result.

The report is tightly constrained and has a lot of ground to cover, so be succinct and to the point. Do not waste words on niceties.

- **Do not** include a cover page, contents list, index or abstract.
- **Do not** include appendices or supplementary materials.
- **Do not** include your own name anywhere in the report, including the filename. Submissions are anonymised, please do not violate that anonymity.

A suggested report structure is given below. This is not a strict prescription – you may organise the material differently if you think doing so presents your work more clearly – but you do need to address all these points.

1. **Data**: Brief introduction to the dataset: what is it, where does it come from, how is it organised?
2. **Strategy**: Explain how you have approached the sonification and how it is structured. How are you mapping your data into sound and why?
3. **Implementation**: Explain the tools and techniques used to realise your sonification. How is the sound generated? What languages or libraries have you used? If your implementation is code-based, you may provide a link to a repository, but do not include any code in your report.
4. **Critical Analysis**: Review the results of your sonification. Identify strengths and weaknesses. Does the audio achieve what you intended?
5. (Optional) **References**: You do not need to provide an extensive bibliography, but you must cite and/or provide links to sources of data, code, tools and inspiration. You can collect those in their own section or include them inline as you go along if it makes more sense to do so.

You may include figures and tables, but they must fit within the page limit. Ensure that all labels and captions are at a legible size – anything we can't read will be treated as not having been included at all.

Marking Criteria

100 marks are available for this coursework, which will contribute 30% of your overall grade for the module. The marks are divided equally between the audio and the report, apportioned as follows:

Audio (50 marks)

- **Aesthetics (10 marks)**
  How interesting or satisfying is the audio to listen to? Note that this does not necessarily imply it should be "pretty", though it can be. But harsh or discordant sounds can also be informative and effective when used appropriately.
- **Technical Merit (20 marks)**
  How well does the implementation function to achieve your aims? Does it manage to translate the data into audio according to the strategy described in the report? How has the sound actually been created? Are the audio parameters appropriate? Is a reasonable volume range maintained? Are the sonic elements well balanced? Does the sound exhibit clipping or aliasing or noise? Is it even **safe** to listen to?
- **Communication & Representation (20 marks)**
  Is the strategy successful? Does the sonification capture or reveal useful features of the data? Is sound a sensible vehicle for this? What does it emphasise and what does it gloss over?

Report (50 marks)

- **Clarity of Explanation (20 marks)**
  Does the report get across what you have done and why? Does it make sense? Can we understand your process and goals, what tools you have used and what sources you are drawing from?
- **Data Analysis & Rationale (15 marks)**
  What do you think might be present in the data and how are you attempting to elucidate it? Do your choices contribute to that goal?
- **Critical Analysis (15 marks)**
  How clearly can you evaluate what you have done? Are you able to identify the strengths and weaknesses, both in the choices made and in the resulting audio? What have you learned from the experience?

It should be clear that there is no single correct answer to this task. We are not expecting you to get it "right", so please don't try. The point of the exercise is to see what choices you make, how you follow through, and what you discover along the way.