Contributors Guidebook for the Data Observatories and Open Collections

Daniel Antal

2023-07-12

Contents

Introduction

note Big data creates **inequalities**, because only the world's biggest global corporations, best endowed universities and strongest governments can maintain long, well-designed, global data collection programs.

This handbook is made for people working together in an open collaboration on the Digital Music Observatory and its funded projects, Listen Local Lithuania, Eviota, SurveyHarmonies and Open Music Europe; the connecting Open Collection Network and the CCSI Data Observatory and Green Deal Data Observatory.

Get in epub or docx formats.

Overview

0.1 Ambition in the Open Music Europe Grant Agreement

Open Music Europe aims to contribute to the aim of creating a decentralised European intelligence hub where centralised data collection and analysis have failed in the music industry during the past 20 years.

We envision the Open Data Observatory as a decentralised, complementary service to the ESSnet system (Eurostat and the national statistics offices) and the planned, centralised European Music Observatory (EMO). [...] We plan to implement the Open Data Observatory as an open-source, public-facing solution working in synergy with the EMO.

We will create `reproducible` "dynamic policy documents" that can be transferred from one country to another or scaled up from one country to a group of countries or even the entire European Union, and recast with little effort after the inception of new data every year, due to their very high level of data/output standardisation and interoperability.

Instead of creating a large budget, ad-hoc valuation reports, these reports, with their automatic data harvesting and processing, can be recast annually with little further investment, or they can be transferred from Hungary to Bulgaria, from Slovakia to Lithuania.

0.2 Methodology in the Open Music Europe Grant Agreement

The `Level 3` version of the *Open Policy Analysis* (`OPA`) Guidelines gives practical guidance on how to improve the transparency and replicability of policy-making work with a scientific underpinning. The OPA guidelines go farther than current Horizon Europe recommendations, subjecting policy research and deliberation to standards as rigorous as those used in e.g. open-source software development and open science peer review.

The guidelines consist of three layers:

- ☒ `open materials` (i.e., the evidence considered in policy) | check out the materials of this handbook.
- ☒ `open analysis` (the analytical procedures to which the evidence is subject); see for example D 1.3 Report on the European Music Economy | Slides | file folder | long term storage | README.
- ☒ `open output` (the indicators, recommendations, etc. derived from the analysis): we syncronize them with libraries via OpenAIRE and our visual assets and cultural data via Europeana aggregators.

Each level must be fully replicable. Establishing a clear link between input and output by displaying how the output changes under `Open Music Europe` uses `Open Policy Analysis` to create policy-relevant indicators. (**??**) and they are fully in line with the Open Science objectives of the European Union (**?**).

Open Music Europe furthermore ensures that all materials, analysis, and outputs meet the `FAIR principles` enshrined in Horizon Europe and the EU's open science agenda and will be submitted as best practices to the *Knowledge4Policy* K4P Portal horizontal knowledge areas on Evidence-Informed Policy Making, Composite Indicators, Modelling, and AI Watch.

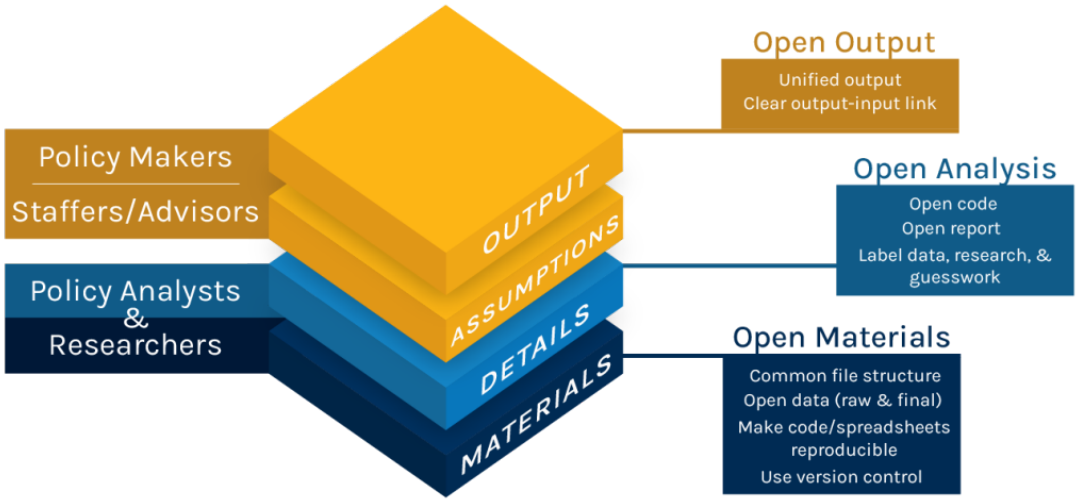

Figure 1: Our ambition to increase transparency with introducing the Open Policy Analysis into the European music policies and collaborative data use in the European industry.

In Europe, the European Commission launched the platform, along with several initiatives to make EU-funded research more accessible, notably, with changing some financial guidelines for its research program, and amending the Open Data Directive in 2019.

The subsequent chapters of this Manual show how we will comply with both FAIR and OPA by making our documents and files Findable in a way that meets open policy analysis compliance, make them Interoperable for machines and researchers, Accessible for human and algorithms, and Reusable for human and automated use.

These efforts toward radical transparency will help ensure open deliberation and consensus-forming among stakeholders. Putting evidence-based policymaking into practice in music research, we will establish a precedent for its incorporation as a keystone of the social sciences and digital humanities more broadly.

0.3 OPA

The OPA Guidelines feature nine steps for implementation, grouped by their relation to the three key principles of OPA: `Open Output`, `Open Analysis`, and `Open Materials`. Each step is modular, meaning it can be implemented independently from the rest, and features three gradual levels of rigor. Analysts can adopt each step at varying levels of rigor, depending on what is appropriate for their work. Status quo in policy analysis practices should be interpreted as Level 0.

You can become a signatory of the `OPA Guidelines` by adding your name (or the name of your organization, if you sign as an organizational representative) to this form and including a statement in policy reports indicating they were developed in accordance with the OPA Guidelines (e.g., "This analysis complies with level 3 of the Open Policy Analysis (OPA) Guidelines developed by BITSS"). Second, you can submit your feedback to these Guidelines to BITSS Project.

0.4 Documents, Word Processing, Manuscripts, Blogposts

We must make our documents Findable, Interoperable, Accessible and Reusable. We introduce the tidy text concept in the Interoperability chapter.

Interoperability in this case making your text editing simpler. Regardless your use of Word, Grammarly, Libre, or a markdown editor, we would like you to keep it as simple that it can be translated to markdown notation, which then can create automatically correct, beautiful Word, PDF, or LaTeX documents.

Markdown is a very simple text notation, similar to the earlier manuscript markup languages used by book editors. You can learn it in 1-2 hours, but even if you do not use it, you should always check that your formatting remains simple enough to be translated (simplified) to markdown. An introduction to using markdown is in **??** Tidy Text. You can get the concept in 10 minutes with Dillinger.

note See the first version of our slide templates here, and visit our template folder (Open Music Europe dissemination and communications assets). Help our colleagues at Synyo to make this template more aesthetic, more usable, and more interoperable by sending issues and big reports here.

0.5 Datasets, Tables, Tabular Data

We must make our documents Findable, Interoperable, Accessible and Reusable. We introduce the tidy data concept in the Interoperability chapter, and go into further depth (only to be used for the Data Management Plan, and some quantiative task of WP1, WP2, WP3, and the programimng parts of WP3) in the **??** Tidy data separate chapter.

For laypersons, tidy data ensures that your data is very easy to import into a relational database (SQL, etc) or it can be read into Excel, OpenOffice, SPSS or STATA without a hiss. Keeping your data is keeping your data simple, and will save you countless hours in the Data Sisyphus.

0.6 Presentations

Presentations are special types of documents where the visual elements dominate. Making them truly interoperable is a big challenge, because we expect them to work on a projector, on a Windows laptop, an Apple desktop, a Linux server, on a tablet, smartphone.

note See the first version of our slide templates here, and visit our template folder (Open Music Europe dissemination and communications assets). Help our colleagues at Synyo to make this template more aesthetic, more usable, and more interoperable by sending issues and big reports here.

0.7 Working Folders (repositories)

Our file folders, or repositories, are stored on Microsoft's Github. They work similarly to a Google Drive or a Dropbox, but with some further featured that makes them compliant with the EU's FAIR and the OPA.

The use of a standardised file structure and the use of an open, selfcontained folder with a readme file is central to **??** Open collaboration, i.e., or project management philosophy, as well as to compliance with the **??** F - Findable principle in the European open science requirements and in OPA.

- ☒ Access to them is not proprietary, it requires the https protocol (or SSH)
- ☒ They are version controlled, and protected for deleting, overwriting
- ☒ Many people can work on them without disturbing each other.
- ☒ WP leaders and task leaders can assign tasks connected to files (such as issues, check out the improvement issues to make our Word and PPTX files truly interoperable).
- ☒ Competing ideas can be disussed under the issues, or help can be asked for, even outside of our work groups or consortia.
- ☐ You need a program to sync it to your computer (Similarly to DropBox o Microsoft's OneDrive)