This is the title

This is the subtitle

Contents

[Data Summary 1](#_Toc140568828)

[Purpose of data generation in relation to the project 3](#_Toc140568829)

[1.2 Definitions 4](#_Toc140568830)

[Size 6](#_Toc140568831)

[1.4.3 Data Summary: Music, Society, Citizenship 6](#_Toc140568832)

[Data generation 7](#_Toc140568833)

[Size 7](#_Toc140568834)

[1.2.4 Data Summary: Innovation 8](#_Toc140568835)

[1.5 Data Dissemination: Observatory 8](#_Toc140568836)

[1.6 Data Reuse 8](#_Toc140568837)

[1. 2. FAIR data 10](#_Toc140568838)

[2.1. Making data findable, including provisions for metadata 10](#_Toc140568839)

[2.1.1 Use of persistent identifiers 10](#_Toc140568840)

[Datasets 11](#_Toc140568841)

[Visualizations. 12](#_Toc140568842)

[Documents 12](#_Toc140568843)

[Metadata 12](#_Toc140568844)

[2.1.2 Rich metadata for discovery 12](#_Toc140568845)

[2.1.3 Search keywords 13](#_Toc140568846)

[2.1.4 Machine actionable metadata 13](#_Toc140568847)

[2.2 Making Data Accessible 13](#_Toc140568848)

[2.2.1 Repositories 13](#_Toc140568849)

[2.2.3 Metadata 14](#_Toc140568850)

[2.3. Making Data Interoperable 15](#_Toc140568851)

[Commonly used ontologies, vocabularies, and data models 16](#_Toc140568852)

[Data exchange and re-use across disciplines 17](#_Toc140568853)

[Data exchange in the music sector 17](#_Toc140568854)

[2.4. Increase data re-use 17](#_Toc140568855)

[Documentation standards 17](#_Toc140568856)

[Good data semantics 18](#_Toc140568857)

[API access 19](#_Toc140568858)

[Data licensing 19](#_Toc140568859)

[Other data 19](#_Toc140568860)

[2.5 Quality Assurance 19](#_Toc140568861)

[Eurostat indicator design principles 19](#_Toc140568862)

[Open source 20](#_Toc140568863)

[Internal peer review 20](#_Toc140568864)

[4. Allocation of resources 20](#_Toc140568865)

[5. Data security 21](#_Toc140568866)

[6. Ethics 21](#_Toc140568867)

[2. Data Catalogue 22](#_Toc140568868)

[References 23](#_Toc140568869)



***Open Music Europe***

***An open, scalable data to-policy pipeline for European music ecosystems***

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Data Summary

Open Music Europe offers a ‘data-to-policy’ pipeline, which extends the music data pipeline to evidence-based business and policy administration. A data pipeline is a method in which raw data is ingested from various data sources and then ported to data store, for further analysis, in this case, to an open, shared, collaborative music observatory. We extend this pipeline using reproducible research techniques, a novel application of the *Open Policy Analysis Guidelines*, and good statistical practices to support evidence-based policy analysis, scientific music research and sound business strategy building. In this last leg of the pipeline, we emphasise usability for our project’s target audiences and good documentation practices. We want to ensure that our data is high quality and well understood to support robust and correct business, scientific or policy conclusions.

The European music sector is fragmented: it is almost exclusively made of micro- and small enterprises and similar-sized organisations that lack in-house data scientists and engineers, statisticians, and often even IT functions. This means that data is scattered in thousands of locations divided by legal and technical barriers, resulting in gaps of actionable data for business, scientific or policy use. The main objective of Open Music Europe is the creation of new statistical indicators that fill the data gaps identified in the *Feasibility Study for the Establishment of a European Music Observatory* (henceforth: European Music Observatory Feasibility Study) (Commission et al. 2020).

Our project follows the organisational structure of this *Feasibility Study* by establishing four data pillars. This Data Management Plan follow the same structure in its *Data Summary* part. Our project work packages follow this structure: WP1—music economy; WP2—diversity and circulation; WP3—societal aspects; WP4—innovative use of data.

We focus on this level of data because it is actionable and can be shared. A statistical indicator summarises a key issue or phenomenon derived from a series of observed facts, a generalisation, which is in most cases void of business confidential or protected personal data. It lets decision-makers focus on the big picture instead of getting lost in data details. Our work packages follow the same structure:

| Objectives | WP1 Economy | WP2 Diversity | WP3 Society | WP4 Innovation |
| --- | --- | --- | --- | --- |
| Objective 1: methodology | T1.1 data needs and methods | T2.1 data needs and methods | T3.1 data needs and methods | T4.1 data needs and methods |
| Objective 2: collection | T1.2 collection and processing | T2.2 collection and processing | T3.2 collection and processing | T4.2 collection and processing |
| Objective 3: empowering | T1.3 reproducible use case | T2.3 collection and processing | T3.3 collection and processing | T4.3 collection and processing |

* **O1 identify European music data gaps**; establishs data sources and methodologies to create valid statistical qualitative and quantitative indicators, or simply indicators (the word statistical means that some statistical procedure is applied to generalise the data.)
* **O2 bridge these gaps** by re-processing or collecting raw data and generating new data assets. T1.2, T2.2, T3.2, T4.2 tasks will acquire (collect or re-use) raw data that can form the basis of business or policy indicators or other research inputs in a processed form.
* **O3 empowers the European music ecosystem’s stakeholders** to use the data to solve important business, research or evidence-based policy problems; place the data into scientific and policy reports (including use cases for business use) to show how the data can be best used, or similar indicators created in other countries, policy contexts or businesses.
* The **Open Music Observatory will disseminate these data assets** so that they are added with new, FAIR features to the Digital Music Observatory minimum viable product (MVP.) In this regard, the projects’ dissemination and data management plans are harmonised. The Digital Music Observatory follows the *Feasibility Study on the European Music Observatory*, and it can be part of a later institutionalised observatory.

Statistical indicators, or, in short, indicators, are derived from accounting systems and statistical data. Accounting systems are coherent and integrated accounts, balance sheets and tables based on a set of agreed rules. Statistical data is also collected based on a set of agreed and coherent rules, for example, with surveys using harmonised and standardised questionnaires to ask music professionals and audiences about their access, participation and perceptions of music.

In Open Music Europe, we use corporate and national accounts to create indicators. Corporate accounts are recorded based on national financial accounting standards, which are greatly harmonised in Europe; the standardisation of sustainability accounts is underway during our projects. The music industry uses special royalty accounts, which collect and organise information about royalties earned, accrued, and paid. National accounts are created with statistical methods by observing all corporate accounts in a country by tax and statistical authorities.

Our data-to-policy pipeline uses the R open-source statistical environment to collect and process data from accounting systems. Our WP4 creates a statistical software ecosystem of the tools used for this purpose. When we work with national accounts data, we reuse and reprocess open statistical information, and the resulting datasets remain open. When we work with corporate accounts data, we will keep our methods, algorithms, and, if possible, code open, but access to the accounting data and its pre-processed form (*accounting microdata*) will be restricted.

The aim of statistical data collection is the creation of sets of data that can be generalised well; for example, when we collect data about people, it includes data on people from all relevant age or gender groups. Open Music Europe is using and developing data collection from primary sources, such as surveys and ephemeral big data, or taking samples of large, pre-existing but biased datasets, such as royalty accounts. The *survey microdata* will be processed into datasets that are void of personal data.

## Purpose of data generation in relation to the project

The Open Music Europe Project itself can be seen as a data management project. In other words, the Data Management Plan is not only aiming at making the data generated or reused by the project further findable, accessible and reusable in an interoperable way, but data generation is a central aim of the grant. Our *Grant Agreement* and the project used the data gap analysis of the *Feasibility study for the establishment of a European Music Observatory*. This data gap analysis is dated and does not contain strictly defined data management instructions or a concise definition. We treat it as a starting point.

The three main objectives of the project are:

* **O1 identify European music data gaps**; T1.1, T2.1, T3.1 and T4.1 will suggest ways how to fill the data gaps identified before Open Music Europe and discovered since and offers stakeholders consultation on them.
* **O2 bridge these gaps** by re-processing or collecting raw data and generating new data assets. T1.2, T2.2, T3.2, T4.2 tasks will acquire (collect or re-use) raw data that can form the basis of business or policy indicators or other research inputs in a processed form.
* **O3 empowers the European music ecosystem’s stakeholders** to use the data to solve important business, research or evidence-based policy problems. This is the role of T1.3, T2.3, T3.3, T4.3.
* The **Open Music Observatory will disseminate these data assets** so that they are added with new, FAIR features to the Digital Music Observatory minimum viable product (MVP.) In this regard, the projects’ dissemination and data management plans are harmonised.

The subsequent subchapters of the **Data Summary** do not present the data in the functional breakup of the Grant Agreement but according to the subdomains or the “pillars” of the EMO Feasibility Study.

* [Data Summary: Music Economy](#pillar1) presents the data management of T1.1 (objective 1), T1.2 (objective 2), T1.3 (objective 3).
* [Data Summary: Diversity and Circulation](#pillar2) presents the data management of T2.1, T2.2, T2.3.
* [1.4.3 Data Summary: Music, Society, Citizenship](#pillar) presents the data management of T3.1, T3.2, T3.3.
* [1.4.4 Data Summary: Innovation](#pillar4) presents the data management of T4.1, T4.2, T4.3.

Apart from this Data Management Plan, or DMP, the project has a Dissemination, Communication and Exploitation Plan (DCE Plan). Because data management is not an auxiliary activity of this project, but the main project objectives are aiming at the generation of new data, in terms of the dissemination and exploitation of data assets and the live policy documents (which empower the stakeholders to use the new data assets functionally), the DCE Plan and the DMP are coordinated.

The main aim of the DMP is to inform potential users of the commonly understood data in the forms of datasets, datacubes, and databases of Open Music Europe visible via metadata and, whenever possible, with data catalogues and disseminated data. As required, we also make provisions for “other data”, i.e. visualisation files, texts, blog posts, and more complex documents in lesser detail, as their access and use require less management information because their dissemination and communication are the focus of the Dissemination, Communication and Exploitation Plan. See [1.5 Data Dissemination: Open Music Observatory](#wp5) for further details.

|  |
| --- |
| Purpose of data generation in relation to the project |
| The most relevant project actions are:   * Develop policy-relevant indicators for the total economic value of music (WP1) * Develop policy-relevant indicators for music diversity and circulation (WP2) * Develop policy-relevant indicators for the societal impact of music, within an sustainable development goal (SDG) framework (WP3) * Develop new survey methods for capturing scarce data * “Provide methodologies for capturing the economic and societal value of music” |

Apart from this, as data management is a primary activity of the project, it has a project action *Develop new software for rendering fragmented, hidden, and unharmonised/unprocessed data usable* that has a supporting role in data collection and data generation.

The Open Music Europe Data Management Plan follows the *Horizon Europe Data Management Plan Template, Version 1* (OpenAire 2021).

## 1.2 Definitions

The following definitions are taken from Article 2 of the *2019/1700 EU Regulation on establishing a common framework for European statistics relating to persons and households, based on data at individual level collected from samples* and the Article 3 of *223/2009 EC Regulation on European statistics* (EUR-Lex 2015; European Parliament and Council of the European Union 2009, 2019). While the latter regulation is no longer in force, its definitions became parts of the European statistical vocabulary.

* statistics means quantitative and qualitative, aggregated and representative information characterising a collective phenomenon in a considered population;
* data collection means surveys and all other methods of deriving information from different sources, including administrative sources;
* microdata means non‐aggregated observations or measurements of characteristics of individual units, without direct identifier;
* administrative records mean data generated by a non‐statistical source, usually a public body, the main aim of which is not the provision of statistics;
* observation unit means an identifiable entity about which data can be obtained, it is also often called a *statistical unit*, which means the basic observation unit, namely a natural person, a household, an economic operator and other undertakings, referred to by the data;
* confidential data means data which allow statistical units to be identified, either directly or indirectly, thereby disclosing individual information. To determine whether a statistical unit is identifiable, account shall be taken of all relevant means that might reasonably be used by a third party to identify the statistical unit;
* use for statistical purposes means the exclusive use for the development and production of statistical results and analyses;
* direct identification means the identification of a statistical (observation) unit from its name or address, or from a publicly accessible identification number;
* indirect identification means the identification of a statistical (observation) unit by any other means than by way of direct identification;

The following definitions are mainly taken from metadata standards, which are described in more detail in the [2. FAIR data](#FAIR-data) of the DMP.

* machine actionable indicates a continuum of possible states wherein a digital object provides increasingly more detailed information to an autonomously-acting, computational data explorer. Our data management practice aims to be machine-actionable to a high degree.
* metadata is a statement about a potentially informative object, usually on how an informative objects such as a dataset or file can be found, accessed, used.
* statistical metadata means information that is needed to use and interpret statistics and that describes data in a structured way;
* datacube is a dataset as a multi-dimensional space, or hyper-cube, indexed by those dimensions.
* dataset is a collection of “collection of data, published or curated by a single agent, and available for access or download in one or more serializations or formats”. the data in a dataset can be described as observations, metadata about the observational structure, structural, and reference metadata. For a more detailed definition, see 2.3.1 (data model and vocabulary.)
* administrative records microdata datasets contain many observations about inventory items or accounts of musical works, biographies, sound recordings. Accounting microdata datasets are not statistically processed and often must be treated as confidential data.
* survey microdata datasets contain observations as responses to a questionnaire by an individual or an enterprise. Survey microdata datasets are not statistically processed and almost always must be treated as confidential data.
* indicator datasets contain statistically processed data, usually statistics of a measured observation, including data manipulations that de-personalize the data.
* visualisations like data visualisations, schematic charts, drawings, photographs, and their collages will be created as still image files.
* data visualisations are still images saved as interoperable files containing visual summaries of data.
* charts are still image files created by their author to visualise logical or temporal relations.
* other visualisations are photographs, drawings or other non-data visualisations in interoperable still image files.
* documents: Documents containing information legally fall under the definition of data. The DMP uses data in the common sense of the word (datasets, datacubes and their visualisations, metadata, databases) and refers to other documents as other data; their FAIRification is included in this document but not their dissemination and communication plan.
* blogpost documents: A special type of document that is intended to be released on the project website and communicated to the public.
* scientific documents: in the context of the DMP, complex documents containing text, visualisations, data tables, bibliographic references as manuscripts, working papers, journal articles and other documents intended for scientific dissemination.
* policy documents: in the context of the DMP, complex documents containing text, visualisations, data tables, and bibliographic references are intended to be used in line with the Open Policy Analysis Guidelines.
* live policy documents: interwoven, complex documents of texts, data, and software code that update policy or business reports as the data sources they are based on are refreshed. They are intended to make the transfer of research to new sites or topics efficient. The Grant Agreement defines D1.2 Report on the European Music Economy, D2.2 Report on Music Diversity and Circulation in Europe; D3.2 Report on Music, Society, and Citizenship in Europe, D4.2 Report on Music Innovation & Technology in Europe as publicly available, exploitable live policy documents.

### Size

The expected size in terms of the row number of this dataset is up to 150,000 observations for Slovakia. Later versions of the DMP will contain data size estimates (or actual sizes) for Bulgaria, Lithuania, and Ukraine. In terms of columns (dimensions, attributes, and measures), we expect that the entire dataset will have at least 30 columns, but not all will be available for unrestricted access.

The dataset is designed to offer foreign keys as links to other datasets in various national databases; this way it will create a linked open data resource that will aim to connect all significant data resources in Slovakia first and then in Bulgaria.

### 1.4.3 Data Summary: Music, Society, Citizenship

|  |
| --- |
| Tip |
| The expected outcome of **Objective 1 MAP the policy and data landscape** is to develop policy-relevant indicators for societal aspects of music. The indicator development guidelines of Eurostat (see *2.5 Quality Assurance*) will be used for the general methodology.  WP3 will analyse and fill data gaps regarding music, society and citizenship in Europe. It will also implement a transferable pilot study to use open data to help MSMEs comply with emerging social and environmental sustainability regulations. The central objectives of WP3 will be to provide indicators that capture social value-adds such as participation in music education, participation in amateur music-making, the role of music in other consumer sectors, and the role of music in social networks; to develop methods for measuring the sustainability of the music industry [*in the context of Sustainable Development Goals*].  Examples of potential indicator candidates developed in WP3 include: Music education—informal practices; Training schemes for music professionals; Social networks and music; Scope of the not-for-profit sector in Europe; Social impact of music in communities; etc. |

The live policy report that will weave together software code, data, legal references and explanation for the reproducible and transferable exploitation of the datasets is deliverable **D3.2 Report on Music, Society, and Citizenship in Europe**. Its open repository contains, from the proposal stage, the document that will be used as hypotheses and inputs to this deliverable in its open policy repository (Open Music Europe 2023b).

For the purposes of environmental impact analysis, national accounts and environmental satellite accounts data will be re-used from open governmental sources.

* T3.2 will create survey micordata datasets following the ESSnet-Culture recommendations on Cultural Access and Participation surveys (Bína, Vladimir et al. 2012; Haan and Adolfsen 2008; Haan and Broek 2012).

Data Reuse

A significant part of the data intended for use in our WP3 and to be disseminated in the Music, Society, Citizenship Pillar is re-used data. The main objective of this work package is to conduct original data collection in a way that is harmonised with pre-existing data, filling the data gaps with optimal synergies. Both the pre-existing data and the new data will be more valuable as they can be linked.

Initially we will reuse data from the following sources:

* The Eurobarometer survey programs microdata databases held by GESIS.
* Microdata of other CAP surveys, pending approval by Eurostat, which were conducted as parts of standardised official statistical surveys, such as AES, by the national statistical offices.
* Survey microdata from the CEEMID project.
* Survey microdata from the SurveyHarmonies project.

### Data generation

The WP3 will generate new data with anonymous and enterprise surveys and integrate them with pre-existing survey metadata.

### Size

The expected size of reusable data cannot be estimated in the 1.0 version of the DMP.

### 1.2.4 Data Summary: Innovation

The live policy report that will weave together software code, data, legal references and explanation for the reproducible and transferable exploitation of the datasets is the deliverable **D4.2 Report on Music Innovation & Technology in Europe as publicly available, exploitable live policy document**. Its open repository contains from the proposal stage the document that will be used as hypotheses and inputs to this deliverable in its open policy repository.

## 1.5 Data Dissemination: Observatory

As a rule of thumb:

1. Music stakeholders in our target group should use the DMP as a data map (metadata) for information about filling their data gaps. The DMP gives an overview of unrestricted and restricted reusable data.
2. Experienced data users or administrators of machine-to-machine data pipelines should consult the DMP on the technical reusability of the data, particularly in reproducible or programmatic data pipelines.
3. Some data assets and some live policy documents contain sensitive information: data that is restricted for some reason, or protected by intellectual property rights. The DCE Plan contains further information on the possibilities for commercial reuse or special, restricted, non-commercial reuse cases.
4. Documents containing information are legally falling under the definition of data. The DMP uses data in the common sense of the word (datasets, datacubes and their visualisations, metadata, databases) and refers to other documents as other data. The dissemination of presentations, journal articles, and other complex documents (that may contain data) is the subject of the *DCE Plan*. The dissemination of methodological documents, including metadata, which makes datasets and datacubes usable for humans or in machine-to-machine communications, can be found in this DMP.

## 1.6 Data Reuse

This is a summary table, for further details, see the following chapters of the DMP.

| Data types | Collecting | Storage | Access |
| --- | --- | --- | --- |
| survey mircodata datasets | Collected by SINUS and REPREX.  Reused under the Open Data Directive. | Sensitive data stored by collecting entity.  Statistical microdata on GitHub repositories (temporary) and Zenodo (long-term) | No access to sensitive microdata.  Non-identifiable datasets will be free to reuse with open data license.  Re-used data will be made further re-usable on similar terms we receive them. |
| administrative records | ALOADED, ARTISJUS, SOZA, MUSICAUTOR and third party administrative records.  Processed by Consortium members. | Sensitive data is stored by collecting entity.  Non-sensitive and limited subsets from M17 in Digital Music Observatory as linked open data. | No access to sensitive microdata.  Limited access to microdata with data linking. Linking is possible with individual licenses. |
| statistical and indicator datasets | Reused from various open government sources.  Processed from Consortium member microdata by Consortium. | Temporary access on GitHub (pre-processed and not human controlled)  Long-term access following monthly release on Zenodo.  From M17 in Digital Music Observatory API. | Released as open data. |
| data visualisations | Created by members of the Consortium or Consortium. | FigShare open science repository. | CC-BY license. |
| software code | UTU, REPREX in open collaboration with qualifying volunteers. | GitHub (development versions)  Comprehensive R Archive Network (CRAN) peer-reviewed releases.  Zenodo (development and peer-reviewed releases.) | Free access, with open source license statement for each software. |
| statistical processing code | UTU, REPREX in collaboration with microdata owners. | GitHub | Whenever possible, released as software.  Redactions are possible for data security purposes. |
| documents | Open Music Europe Consortium and its members | GitHub (see references)  Final versions on Zenodo (and in journals) | Free access for non-commercial use. |
| blogpost documents | Open Music Europe Consortium and its members | On the Open Music Europe project website and on the Digital Music Observatory web resources’ website interface. | CC-BY license |
| live policy documents | Open Music Europe Consortium and its members | Digital Music Observatory web resource | Complex document, parts (data, visualisation, text, code) under various CC licenses and open source licenses. |
| metadata | Open Music Europe Consortium and its members | See details of DMP. | CC0 |

# 2. FAIR data

## 2.1. Making data findable, including provisions for metadata

|  |
| --- |
| Making data findable |
| * Will data be identified by a persistent identifier? * Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. * Will metadata be offered in such a way that it can be harvested and indexed? |

## 2.1.1 Use of persistent identifiers

|  |
| --- |
| Making data findable |
| * Will data be identified by a persistent identifier? |

Our data collection, processing and incremental addition to our datasets will be a daily activity, and we will use temporary (buffer) and persistent identifiers for all data assets.

ORCiD: ORCID provides a persistent digital identifier (an ORCID iD) that you own and control, and that distinguishes you from every other researcher. We use it to identify contributors to our data assets.

DOI: All our datasets will have a versioned DOI on Zenodo; high-frequency data will have temporary identifiers and will only be periodically released on Zenodo with a new DOI. All our visualizations will have a DOI on FigShare, which is a global open repository particularly designed to make visualizations re-usable. When datasets are visualised, then the visualization will be connected to the persistent identifier (DOI) of the dataset.

ISMN: ISMN is the International Standard Music Number. It is a unique number for the identification of all notated music editions from all over the world. We use it in the creation of diversity indicators, in the original inventory microdata datasets.

ISNI: ISNI is the ISO certified global standard number for identifying the millions of contributors to creative works and those active in their distribution. We use it in the creation of diversity indicators, in the original inventory microdata datasets.

ISRC: The International Standard Recording Code (ISRC) is an international standard code for uniquely identifying sound recordings and music video recordings. We use it in the creation of diversity indicators, in the original inventory microdata datasets.

ISWC: The ISWC (International Standard Musical Work Code) is a unique, permanent and internationally recognized reference number for the identification of musical works. We use it in the creation of diversity indicators, in the original inventory microdata datasets.

VIAF: VIAF explores virtually combining the name authority files of national level authority files into a single name authority service. We use the VIAF ID as a PID. We use it in the creation of diversity indicators, in the original inventory microdata datasets.

### Datasets

The datasets of Open Music Europe will be placed in a data catalogue that conforms the [Data Catalog Vocabulary (DCAT) - Version 2](https://www.w3.org/TR/vocab-dcat-2/) W3C Recommendation (W3C 2020). Each dataset will received a versioned DOI as a PID.

DCAT enables a publisher to describe datasets and data services in a catalog using a standard model and vocabulary that facilitates the consumption and aggregation of metadata from multiple catalogs. This can increase the discoverability of datasets and data services. It also makes it possible to have a decentralized approach to publishing data catalogs and makes federated search for datasets across catalogs in multiple sites possible using the same query mechanism and structure. Aggregated DCAT metadata can serve as a manifest file as part of the digital preservation process.

### Visualizations.

|  |
| --- |
| PID for Visualizations: All our visualisations will have a DOI on FigShare |

### Documents

Live policy documents: Live policy documents, as their name suggests, are live updating, often very frequently. Similarly to their (embedded) high-frequency datasets, they will be saved periodically and released on Zenodo with a versioned DOI.

Static documents: Under the current European legislation, all informative files, including text documents fall under the legal definition of data. Their dissemination is defined in the Open Music Europe consortium’s Dissemination, Communication and Exploitation Plan. The DMP relates to the human and machine-actionable release of more structure information that is commonly referred to as ‘data’. To make our data more reusable, all data used in the disseminated static documents will be made available in human and machine-actionable dataset forms and data visualisations.

### Metadata

### 2.1.2 Rich metadata for discovery

|  |
| --- |
| Making data findable |
| * Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. |

All our data will be provided with the mandatory metadata of the *DataCite 4.4* standard (DataCite Metadata Working Group 2021), and in almost all cases, all recommended metadata, too.

### 2.1.3 Search keywords

|  |
| --- |
| Note |
| * Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use? |

Subject keywords will be used from global library services to optimize the possibility of discover and re-use. We will use three subject-heading lists:

* We mainly use the Library of Congress Subject Headings ([LCSH](https://www.loc.gov/aba/cataloging/subject/)), which is perhaps the most widely adopted subject indexing language in the world, has been translated into many languages, and is used around the world by libraries large and small. LCSH has been actively maintained since 1898 to catalog materials held at the Library of Congress.
* When the LCSH, which is a general-purpose vocabulary, is not granular enough, we will use the [THESOZ Thesaurus](https://lod.gesis.org/thesoz/en/) and [STW Thesaurus for Economics](https://lod.gesis.org/stw/en/) as multi-language vocabularies that are particularly useful for making our surveys well harmonised.
* For increased usability, when possible, we will connect these keywords (subjects) to [Wikidata Properties](https://www.wikidata.org/wiki/Q89560413).

## 2.1.4 Machine actionable metadata

|  |
| --- |
| Note |
| * Will metadata be offered in such a way that it can be harvested and indexed? |

Our data management practice is machine-actionable, this means that we will offer metadata in a way that it can be machine-read, harvested and indexed.

## 2.2 Making Data Accessible

Open Music Europe, as the name of the project suggests, want to make as much data openly accessible for music businesses, music-related and evidence-based cultural policy and music research as possible. Most of our data in the form of datasets, and other data (i.e. visualizations, texts, complex documents and live policy documents) will be openly accessible, unless otherwise stated in this DMP.

## 2.2.1 Repositories

|  |
| --- |
| Note |
| * Will the data be deposited in a trusted repository? * Have you explored appropriate arrangements with the identified repository where your data will be deposited? * Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object? |

Our main, long-term repository for data is Zenodo. The project’s 5.1 deliverable, the Open Music Observatory, had a minimum viable product prototype, the Digital Music Observatory, which has had a community on Zenodo since 2021. (Digital Music Observatory 2021b)

Our short-term repository, following the Open Policy Analysis Guidelines, provides access and reviewability to our data and inputs from the very early stage (in some cases, regarding documentation, from the proposal phase.) These repositories are stored on GitHub, and their addresses are available in the reference list of this DMP and via our project website.

Zenodo and GitHub offer continuous integration with each other. GitHub will be used as a working repository to comply with the OPA Level 3 standard, and data assets that are approved for dissemination and communication will be placed onto Zenodo.

FigShare is a repository aimed at reusable visualisations. While we can also place visualisations on GitHub and Zenodo, FigShare has a visualization-centric global audience. When visualisations can be reused individually, we will place them FigShare, and use FigShare’s DOIs as persistent identifiers. Naturally, when the primary DOI and dissemination point of a dataset and its methodological description is Zenodo and the visualisations’ is FigShare, we will cross-reference these assets. The project’s 5.1 deliverable, the Open Music Observatory, had a minimum viable product prototype, the Digital Music Observatory, which has had a community on Zenodo since 2021. (Digital Music Observatory 2021a)

### 2.2.3 Metadata

|  |
| --- |
| Metadata |
| * Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data? * How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available? * Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)? |

The metadata of Open Music Europe data will be openly available and in human-readable, re-usable, and machine-readable format(s).

In the first phase of the project (M1-M16), this will be achieved with periodically releasing all data on the Zenodo open science repository, in the Digital Music Observatory collection. All data deposited and released in the repostitory will have metadata available in the following metadata formats: MARCXML, Dublin Core (according to OpenAIRE Guidelines), DataCite, DCAT, JSON-LD (Schema.org).

In the second phase of the project, we will start delivering the D5.1 Open Data Observatory, which will provide API access to our data (via SQL querries and in machine-readable JSON format) and provide further, machine readable data catalogues along the aforementioned solution.

The metadata (and the data) will be deposited on Zenodo, which provides access to the data for the lifetime of the host laboratory CERN, which currently has an experimental programme defined for the next 20 years at least.

According to the *Exploitation Plan* of the Open Music Europe Consortium, our plan is to make our Open Music Observatory, as part of a larger Digital Music Observatory eventually part of the planned European Music Observatory. If this is not feasible, we would like to find an exploitation path that ensures the continuity of the data collection, processing, and dissemination of the project. If successful, then we will not only provide long-term access to the data and the metadata but continue to renew our datasets.

Almost all data will be provided in a way that it can be used and re-used in numerous, widely used open sources and licensed software applications. To enhance interoperability, we will always provide data CSV files that conform the W3C interoperability standard about the release of CSV files with metadata.

After consulting stakeholders in the target group of our project, we may provide data in different file formats that is more comfortable for our users.

One of our project aims is to provide a fully reproducible data-to-policy pipeline and full reproducibility in the R statistical environment and language. The R environment is both open-source and interoperable, it is available on a wide range of operational systems (i.e., various Linux or BSD distirbutions, including MacOS, and various Windows operational systems.) Apart from using the resilient but inefficient CSV format, we will release data in native rdf formats. One of our aims in our Work Package 4 is to create an extension package to the R language that ensures the

## 2.3. Making Data Interoperable

|  |
| --- |
| Note |
| * What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones? * In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them? * Will your data include qualified references to other data (e.g. other data from your project, or datasets from previous research)?   A qualified reference is a cross-reference that explains its intent. For example, X is regulator of Y is a much more qualified reference than X is associated with Y, or X see also Y. The goal therefore is to create as many meaningful links as possible between (meta)data resources to enrich the contextual knowledge about the data. (Source: https://www.go-fair.org/fair-principles/i3-metadata-include-qualified-references-metadata/) |

As data interoperability intersects with data re-use, we explain some of our approach in 2.4.

### Commonly used ontologies, vocabularies, and data models

Our datasets follow the vocabularies and definitions of the data model of the Statistical Data and Metadata Exchange (SDMX), i.e., the SDMX 3.0 Technical Specifications. Section 2 Information Model: UML Conceptual Design. Version 3.0 which provides the widest possible interoperability with the data products of national statistical agencies and international agencies such as Eurostat or OECD (SDMX 2021). We will also comply with the W3C’s RDF Data Cube Vocabulary, which is the application of the datacbue and datsaset model of SDMX to word wide web (W3C 2014).

These definitions were designed for statistical datasets, and they are directly applicable to our indicator datasets. However, with simplifications, they can also apply to our microdata, except that these datasets contain statistically not yet processed raw data. This means that the microdata datasets will contain the measures and dimensions as grouping variables, which will become the measures and dimensions of the indicator dataset after statistical aggregation.`

In T3.2 we will aim to apply more and more of the DDI Standards, starting with the DDI-Codebook 2.5 standard (DDI Alliance 2012). Later versions of this DMP will provide guidance on the introduction of further DDI standards.

Our music diversity and circulation datasets will follow the Polifonia Ontology Network ontologies (Berardinis et al. 2023), curated and developed by the *Polifonia: a digital harmoniser for musical heritage knowledge* Project (Polifonia 2021). We consider these ontologies as the best available ontologies concerning metadata about sound recordings, works, performances and musicological variables. Using these ontologies will be introduced in the indicator and dataset development of T2.1 and T2.2 and further reference will be added to this DMP in later versions.

Qualified references

|  |
| --- |
| Note |
| This part is basically a rewording of the requirements with the exception of handling the issue of ISWC that needs to be clarified with CISAC. |

If the data set builds on another data set, if additional data sets are needed to complete the data, or if complementary information is stored in a different data set, this needs to be specified. This will be the case in many of our datasets.

From a data documentation point of view, a scientific link between the data sets needs will be described, and all data sets will be cited with the inclusion of their persistent identifiers.

This approach will be taken with our indicator datasets, which follow as closely as possible SDMX practices; complex indicators will often use as components of pre-existing datasets disseminated following the SDMX standards.

### Data exchange and re-use across disciplines

As data interoperability intersects with data re-use, we explain our interdisciplinary approach in 2.4.x. We mainly foster interdisciplinary re-use of our indicator datasets by applying the cross-domain codebooks for data attributes and dimensions of SDMX, which are precisely designed for cross-domain interoperability. Likewise, we will foster interdisciplinary re-use of our microdata datasets by applying the DDI Codebook standard and THESOZ interdisciplinary, multi-language concept mapping to all our variables.

### Data exchange in the music sector

Apart from using the Polifonia Ontology Network ontologies, which we regard as the highest standard of music sector interoperability, we will use several international standards.

ISRC: Whenever possible, we will identify sound recordings with their standard, unique, international identifier.

ISWC: The International Standard Musical Work Code (ISWC) identifies musical works as intangible creations. In the case of the diversity and circulation datasets, the use of ISWC codes as internationally harmonised PIDs would be desirable; however, by practice, CISAC currently discourages the mass publication of ISWCs. In the pilot phase of T2.3, when we create the pilot of the SCMDb, we will address CISAC directly for best policies. As a workaround, we will create potential identifiers that can remain persistent identifiers in Slovakia.

## 2.4. Increase data re-use

Our Open Music Europe project aims to create a model for creating reusable data for the European music sector. We take several measures to increase reusability, which partly overlap with interoperability measures.

Our data management makes several provisions that improve machine-to-machine (machine actionable) and human reuse of the data.

### Documentation standards

|  |
| --- |
| Note |
| * How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)? * Will the provenance of the data be thoroughly documented using the appropriate standards?   OPA (Open Analysis) 3: Code is clearly documented into a dynamic document, or open notebook. No spreadsheets.  OPA (Open Materials) 6: All project components are organized in a selfcontained folder using a Standard File Structure (SFS), and a readme file is included. |

All our datasets are created in compliance with the OPA Guidelines (Level 3) standard, which requires that the data is presented with clear documentation and with the code that produced the dataset. The selfcontained folders with SFS, README can be found in the reference list of the DMP and via the Open Music Europe project website.

The data provenance will be documented according to the general standards of OPA Guidelines, and more specifically by the provenance documentation standards of the data in question. In case of survey data, we will follow DDI standards, and in the case of statistical data, SDMX standards and practices. In case of music diversity and circulation data, we will follow the Polifonia Ontology Network’s provenance documentation standards. These are explained in more detail earlier.

In the second phase of the project, D5.1 Open Music Observatory will provide an improved human- and machine-readable data catalogue. The DMP will be updated accordingly.

### Good data semantics

Data where the structure of the data reflects well the meaning of the data is particularly easy to be reused by humans.

Our datasets will always follow the **tidy data principle**, which is a statistical reformulation of Codd’s 3rd normal form (3NF), an important and basic principle of database design. Tidy data usually requires little or no further processing before use in spreadsheet or statistical or other software applications, and the in the case of smaller datasets, an analyst can usually read the data.

The datasets follow the definitions of SDMX and W3C, and organized in a way that is expected by machine-to-machine communication and by experienced data users. They follow the datacube model, where the columns are organized as measures, dimensions and attributes. Each column has has a standardised name, which follows the snake\_case convention with lowercase ASCII characters for easier programmatic reuse, and interoperability.

The observations are provided with unique identifiers, and whenever possible, with PID identifiers for full interoperability.

The equivalence of a good data semantics in the case of visualizations is the use of visual forms that generally not misleading for a human user.

The equivalence of good data semantics in the case of text and complex documents is a clear structure which includes a summary and a full list of references.

### API access

The Open Music Europe project foresees the provision of API access to our datasets, with exceptions laid out in later versions of the DMP, after Month 17 of the project. The API access will mainly foster machine-to-machine data reuse, and professional users who prefer to query data with the SQL or ddplyr data querying languages.

### Data licensing

|  |
| --- |
| Note |
| * Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement? * Will the data produced in the project be useable by third parties, in particular after the end of the project? |

Our project, as we have outlined in the [Data summary](#data-summary), is unusual because, in our case, data generation and data reuse is not a secondary activity of the project, but relate to its main objectives. In this respect, our project is a data innovation project.

This means that the project will use much project background, which is often IP-protected data, and will create, in some cases, a project foreground that is IP protected. We also foresee the reuse of important data that is not unrestricted in reuse or completely restricted to be used without a license.

As a rule of thumb, whenever we reuse open data, or we engage in primary data collection and generation as a granted activity, we will create data that can be reused with a CC0 or similar license if it is statistically aggregated. We will also make subsets of our microdata available for replication.

Regarding re-processed, IP-protected data, and microdata, potential reuse conditions will be explored in more detail in the Exploitation Plan of the project.

### Other data

|  |
| --- |
| Note |
| * Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects. |

## 2.5 Quality Assurance

|  |
| --- |
| Note |
| Describe all relevant data quality assurance processes. |

### Eurostat indicator design principles

In our projects, we follow the best practices of key business information, statistical, and evidence-based policy indicator design. In doing so, we would like to find synergies among various recent innovations in statistics and open science. Throughout the project, we will follow the Eurostat guidelines on creating new indicators (Eurostat 2014, 2017; Kotzeva et al. 2017), which will ensure broad consensus-forming among stakeholders around the objectives and methodology of the improved measurements.

### Open source

### Internal peer review

Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects

1. Other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made

## 4. Allocation of resources

|  |
| --- |
| Note |
| What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe |
| Note |
| Who will be responsible for data management in your project? |

* Within the Open Music Europe consortium, ever organization is responsible for the management of datasets that are created for the delivery of their tasks.
* REPREX is responsible for maintaining the DMP as a living document and maintaining the Data Catalogue. The task is lead by Daniel Antal.
* REPREX is responsible for setting up the Open Music Observatory on the basis of the [Digital Music Observatory](https://music.dataobservatory.eu/) minimum viable product and create a rich documentation and various access modes for our data.
* According the Dissemination, Communication and Exploitation Plan of the Consortium, the *Open Music Observatory* will be a part of the future European Music Observatory, or an ongoing, long-term Digital Music Observatory. This means that according to our plans not only the dataset will be preserved after the conclusion of Open Music Europe as a project, but the datasets will contain new observations and corrections beyond the project timeline.

|  |
| --- |
| Note |
| How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who |

Zenodo is hosted by CERN which has existed since 1954 and currently has an experimental programme defined for the next 20+ years. CERN is a memory institution for High Energy Physics and renowned for its pioneering work in Open Access. Organisationally Zenodo is embedded in the IT Department, Collaboration Devices and Applications Group, Digital Repositories Section (IT-CDA-DR).

## 5. Data security

|  |
| --- |
| Note |
| * What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)? * Will the data be safely stored in trusted repositories for long term preservation and curation? |

As our project is a data management focused project, almost all Consortium members will store, archive or transfer sensitive data. As outlined in 1. [Data summary](#data-summary) we follow the *2019/1700 EU Regulation on establishing a common framework for European statistics relating to persons and households, based on data at individual level collected from samples*, and we expect Consortium members to continue to perform these duties in a competent manner, and provide microdata or statistical data in a form that disallow identification.

Our data-to-policy pipeline, following the Open Policy Analysis Guidelines uses secure and public repositories for temporary storage or the storage of pre-processed data. These repositories are stored on GitHub (see References), often with so-called continuous integration, which means that every time or code collects or processes data, it will immediately save it to these common file folders.

GitHub is GDPR compliant, and places on a great emphasis on trans-Atlantic compliance, too. GitHub offers AICPA System and Organization Controls (SOC) 1 Type 2 and SOC 2 Type 2 reports with IAASB International Standards on Assurance Engagements, ISAE 2000, and ISAE 3402 for GitHub Enterprise Cloud.

The final, processed data will be synchronized between GitHub and Zenodo, or permanent, long-term repository monthly. Our datasets are periodically updated on Zenodo (under a versioned DOI) which is planned to operate at least for the next 20+ years. Zenodo itself is funded by OpenAIRE and offers continuous integration with GitHub.

## 6. Ethics

|  |
| --- |
| Note |
| * Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA). * Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data? |

Consortium members and their subcontractors will be involved in creating survey metadata datasets and follow the ICC/ESOMAR International Code on Market, Opinion and Social Research and Data Analytics sets the standard of ethical and professional conduct for the global data, research and insights community.

In the case of personal surveys, the rules of GDPR and ICC/ESOMAR will be used to prevent the leakage of personal data into the survey microdata sets, which will be available for analysis. In the next processing step, indicator datasets will be created, which are statistically aggregated from a significant number of responses so that they cannot be disaggregated to guess, estimate, or predict individual personal data.

In the case of enterprise surveys, the survey is designed in a way that it applies only to legal persons. In this case, the provisions similar to the *2019/1700 EU Regulation on establishing a common framework for European statistics relating to persons and households, based on data at individual level collected from samples*, considering, of course, the difference that we do not apply the rules as a statistical authority and have different legal basis for data collection.

# Data Catalogue

DCAT is an RDF vocabulary designed to facilitate interoperability between data catalogs published on the Web. This document defines the schema and provides examples for its use.

The data catalogue will contain in human readable tables and machine readable (DCAT Version 2 standard) format information about the datasets created by Open Music Europe. See an example from the standard:

| Properties | Example values |
| --- | --- |
| dct:title | "Open Music Europe Example Dataaset"@en |
| dcat:keyword | "music"@en, "musique"@fr, "payments"@en |
| dct:creator | Jane Doe |
| dct:isseud | “2021-06-01”^^xsd:date |
| dct:temporal | <http://reference.data.gov.uk/id/quarter/2006-Q1> |
| dcat:temporalResolution | “P1D”^^xsd:duration |
| dct:spatial | <http://sws.geonames.org/6695072/> |
| dct:publisher | open-music-europe |
| dct:language | <http://id.loc.gov/vocabulary/iso639-1/en> |
| dct:accrualPeriodicity | <http://purl.org/linked-data/sdmx/2009/code#freq-A> |
| dcat:distribution | dataset-001-csv |

dataset-001 a dcat:Dataset

## References

Berardinis, Jacopo de, Andrea Poltronieri, Nicolas Lazzari, and Valentina Carriero. 2023. “Polifonia Ontology Network (PON).” Zenodo. <https://doi.org/10.5281/zenodo.7919970>.

Bína, Vladimir, Chantepie, Philippe, Deboin, Valérie, Kommel, Kutt, Kotynek, Josef, and Robin, Philippe. 2012. “ESSnet-CULTURE, European Statistical System Network on Culture. Final Report.” Edited by Frank, Guy. <https://ec.europa.eu/assets/eac/culture/library/reports/ess-net-report_en.pdf>.

BITSS. 2019. “Guidelines for Open Policy Analysis.” Berkeley Initiative for Transparency in the Social Sciences. <http://www.bitss.org/wp-content/uploads/2019/03/OPA-Guidelines.pdf>.

Commission, European, Sport Directorate-General for Education Youth, Culture, M Clarke, P Vroonhof, J Snijders, A Le Gall, et al. 2020. *Feasibility Study for the Establishment of a European Music Observatory : Final Report*. Publications Office. <https://doi.org/doi/10.2766/9691>.

DataCite Metadata Working Group. 2021. “DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs. [Version 4.4.].” DataCite e.V. <https://doi.org/10.14454/3w3z-sa82>.

DDI Alliance. 2012. “DDI-Codebook 2.5 Data Documentation Initiative.” DDI Alliance. <https://ddialliance.org/Specification/DDI-Codebook/2.5/>.

Digital Music Observatory. 2021a. “Digital Music Observatory Collection on Figshare.” Digital Music Observatory. <https://figshare.com/account/home#/collections/5784023>.

———. 2021b. “Digital Music Observatory Community on Zenodo.” Digital Music Observatory. <https://zenodo.org/communities/music_observatory/>.

EUR-Lex. 2015. *Consolidated Text: Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European Statistics and Repealing Regulation (EC, Euratom) No 1101/2008 of the European Parliament and of the Council on the Transmission of Data Subject to Statistical Confidentiality to the Statistical Office of the European Communities, Council Regulation (EC) No 322/97 on Community Statistics, and Council Decision 89/382/EEC, Euratom Establishing a Committee on the Statistical Programmes of the European Communities (Text with Relevance for the EEA and for Switzerland)*. <http://data.europa.eu/eli/reg/2009/223/2015-06-08/eng>.

European Commission. 2023. “Horizon Europe (HORIZON) Programme Guide. Version 3.0.” European Commission. <https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf>.

European Parliament and Council of the European Union. 2009. *Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European Statistics and Repealing Regulation (EC, Euratom) No 1101/2008 of the European Parliament and of the Council on the Transmission of Data Subject to Statistical Confidentiality to the Statistical Office of the European Communities, Council Regulation (EC) No 322/97 on Community Statistics, and Council Decision 89/382/EEC, Euratom Establishing a Committee on the Statistical Programmes of the European Communities (Text with Relevance for the EEA and for Switzerland)*. <http://data.europa.eu/eli/reg/2009/223/oj/eng>.

———. 2019. “Regulation (EU) 2019/1700 of the European Parliament and of the Council of 10 October 2019 Establishing a Common Framework for European Statistics Relating to Persons and Households, Based on Data at Individual Level Collected from Samples, Amending Regulations (EC) No 808/2004, (EC) No 452/2008 and (EC) No 1338/2008 of the European Parliament and of the Council, and Repealing Regulation (EC) No 1177/2003 of the European Parliament and of the Council and Council Regulation (EC) No 577/98 (Text with EEA Relevance).” <http://data.europa.eu/eli/reg/2019/1700/oj/eng>.

Eurostat. 2014. *Towards a Harmonised Methodology for Statistical Indicators — Part 1: Indicator Typologies and Terminologies*. 2014th ed. Vol. 1. Towards a Harmonised Methodology for Statistical Indicators 1. Luxembourg: Publications Office of the European Union. <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-14-011>.

———. 2017. *Towards a Harmonised Methodology for Statistical Indicators — Part 3: Relevance for Policy Making*. 2017th ed. Vol. 3. Towards a Harmonised Methodology for Statistical Indicators 1. Luxembourg: Publications Office of the European Union. <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-17-007>.

Haan, Jos de, and Anna Adolfsen. 2008. *De Virtuele Cultuurbezoeker - Publieke Belangstelling Voor Cultuurwebsites*. SCP-Publicatie 2008/9. Den Haag, the Netherlands: Sociaal en Cultureel Planbureau. <https://archief18.archiefweb.eu/archives/archiefweb/20200311100055/https://www.scp.nl/dsresource?objectid=5c6903f1-e6ae-4e62-9a9b-6d5e1529756a&type=org>.

Haan, Jos de, and Andries van den Broek. 2012. “Nowadays Cultural Participation - an Update of What to Look for and Where to Look for It.” In *ESSnet-CULTURE, European Statistical System Network on Culture. Final Report.*, 397–417. Luxembourg. <https://ec.europa.eu/assets/eac/culture/library/reports/ess-net-report_en.pdf>.

Hoces de la Guardia, Fernando, Sean Grant, and Edward Miguel. 2020. “A framework for open policy analysis.” *Science and Public Policy* 48 (2): 154–63. <https://doi.org/10.1093/scipol/scaa067>.

Kotzeva, Mariana, Anton Steurer, Nicola Massarelli, and Mariana Popova, eds. 2017. *Towards a Harmonised Methodology for Statistical Indicators — Part 2: Communicating Through Indicators*. 2017th ed. Vol. 2. Towards a Harmonised Methodology for Statistical Indicators 1. Luxembourg: Publications Office of the European Union. <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-17-001>.

Open Music Europe. 2023a. “Repository: Report on Music Diversity and Circulation in Europe. GitHub.” Open Policy Analysis Repository with Standardized File Structure. 2023. <https://github.com/dataobservatory-eu/report-music-diversity-circulation-europe/>.

———. 2023b. “Repository: Report on Music, Society, and Citizenship in Europe. GitHub.” Open Policy Analysis Repository with Standardized File Structure. March 25, 2023. <https://github.com/dataobservatory-eu/report-music-society-citizenship>.

———. 2023c. “Repository: Report on the European Music Economy. GitHub.” Open Policy Analysis Repository with Standardized File Structure. March 26, 2023. <https://github.com/dataobservatory-eu/report-european-music-economy>.

OpenAire. 2021. “Horizon Europe Data Management Plan Template, Version 1.” <https://www.openaire.eu/images/Guides/HORIZON_EUROPE_Data-Management-Plan-Template.pdf>.

OpenMuse. 2023. “Open Music Europe (OPENMUSE) – an Open, Scalable, Data-to-Policy Pipeline for European Music Ecosystems.” <https://doi.org/10.3030/101095295>.

Polifonia. 2021. “Polifonia: A Digital Harmoniser for Musical Heritage Knowledge.” CORDIS Community Research; Development Information Service. <https://doi.org/10.3030/101004746>.

SDMX. 2021. “SDMX 3.0 Technical Specifications. Section 2 Information Model: UML Conceptual Design. Version 3.0.” Statistical Data; Metadata eXchange. <https://sdmx.org/wp-content/uploads/SDMX_3-0-0_SECTION_2_FINAL-1_0.pdf>.

W3C. 2014. “RDF Data Cube Vocabulary.” Edited by Richard Cyganiak and Dave Reynolds. W3C. <https://www.w3.org/TR/vocab-data-cube/>.

———. 2020. “Data Catalog Vocabulary (DCAT) - Version 2.” Edited by David AND Cox Albertoni Riccardo AND Browning. W3C. <https://www.w3.org/TR/vocab-dcat-2/>.