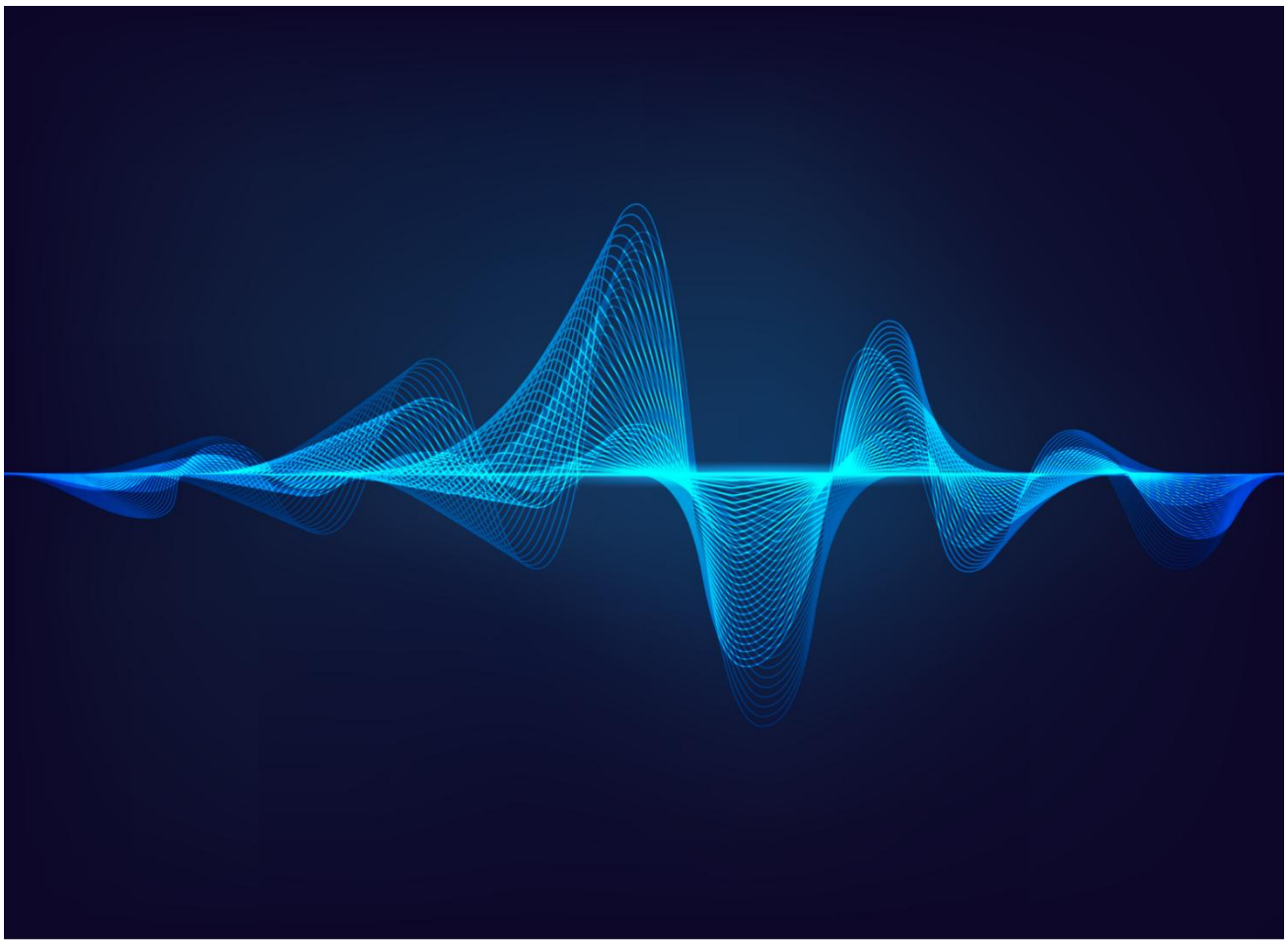


D1.4 – Methodology report on industry partners

OpenMusE

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

Funded by the European Union

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.

Project Factsheet

Acronym:	**OpenMusE**
Title:	**An open, scalable data to-policy pipeline for European music ecosystems**
Coordinator:	SINUS Markt- und Sozialforschung GmbH
Reference:	101095295
Type:	Research and Innovation Action
Program:	Horizon Europe
Start:	1st January 2023
Duration:	36 months
Website:	https://www.openmuse.eu/
Consortium:	**SINUS Markt- und Sozialforschung GmbH,** Germany (SINUS), Coordinator **TURUN YLIOPISTO,** Finland (UTU) **UNIVERSITEIT VAN AMSTERDAM,** Netherlands (UVA) **Scuola Superiore di Studi Universitari e di Perfezionament,** Italy (SSSA) **EKONOMICKA UNIVERZITA V BRATISLAVE,** Slovakia (EUBA) **Reprex B.V.,** Netherlands (REPREX) **SYNYO GmbH,** Austria (SYNYO) **MUSIC INNOVATION HUB SPA IMPRESA SOCIALE,** Italy (MIH) **Slovenský ochranný zväz autorský pre práva k hudobným,** Slovakia (SOZA) **Aloaded AB,** Sweden (ALOADED) **Music Export Ukraine,** Ukraine (MEU) **Muzikos Eksporto Fondas,** Lithuania (MXF) **ARTISJUS MAGYAR SZERZOI JOGVEDO IRODA EGYESULET,** Hungary (ARTISJUS) **MUSICAUTOR SDRUZHENIE,** Bulgaria (MUSICAUTOR) **HEARDIS! GMBH,** Germany (HEARDIS)

© 2023 OpenMusE | HORIZON-CL2-2022-HERITAGE-01-05 | Grant Agreement No. 101095295

Deliverable Factsheet

Number:	**D1.4**
Title:	**Methodology report on industry partners**
Lead beneficiary:	REPREX
Work package:	WP1
Task:	T1.4
Dissemination level:	Public
Submission date:	30.07.2025
Revision/resubmission:	15.05.2026
Main author(s):	Daniel Antal, Reprex , Anna Márta Mester, Reprex
Contributor(s):	Richard Demčák, SOZA - collective management interoperability, Matej Grochal Wikimedia Slovensko, Music Centre Slovakia - Wikipedian in residence, Marián Jankovič, SOZA - legal interoperability, Anna Žilková, Music Center Slovakia - library interoperability
Quality check:	Joost Poort, UvA, Arianna Martinelli, SSSA

Document history:

https://github.com/dataobservatory-eu/music-economy-register/commits/main/

Revision	Date	Main modification	Author
0.1	20/04/2025	First draft finalized	Daniel Antal, Anna Mester
0.2	22/04/2025	Quality Check 1	Joost Poort
0.3	27/04/2025	Quality Check 2	Arianna Martinelli
1.0	29/04/2025	Feedback added	Daniel Antal, Anna Mester
1.1	24/07/2025	Deliverable formatted and finalized for upload	J. Rhys Edwards, Isabella Tautscher, Lisa Littmann
2.0	07/05/2026	Revision based on Periodic review comments	Daniel Antal

© 2023 OpenMusE | HORIZON-CL2-2022-HERITAGE-01-05 | Grant Agreement No. 101095295

Disclaimer of Warranties

This project has received funding from the European Union's Horizon Europe, research and innovation programme, under Grant Agreement No. 101095295.

This document has been prepared by OpenMusE project partners as an account of work carried out within the framework of the EC-GA contract no 101095295.

Any dissemination of results must indicate that it reflects only the author's view and that the Commission Agency is not responsible for any use that may be made of the information it contains.

Neither Project Coordinator, nor any signatory party of OpenMusE Project Consortium Agreement, nor any person acting on behalf of any of them:

(a) makes any warranty or representation whatsoever, express or implied,
  (i). with respect to the use of any information, apparatus, method, process, or similar item disclosed in this document, including merchantability and fitness for a particular purpose, or
  (ii). that such use does not infringe on or interfere with privately owned rights, including any party's intellectual property, or
  (iii). that this document is suitable to any particular user's circumstance; or

(b) assumes responsibility for any damages or other liability whatsoever (including any consequential damages, even if Project Coordinator or any representative of a signatory party of the OpenMusE Project Consortium Agreement, has been advised of the possibility of such damages) resulting from your selection or use of this document or any information, apparatus, method, process, or similar item disclosed in this document.

Glossary

Registers and their records

register: A register aims to be a complete list of the objects in a specific group of objects or population." (Anders and Britt 2007); is a dataset or database that systematically maintains and updates information on a specific set of entities (e.g., individuals, organisations, or events), often with legal or statistical purpose. In statistics, a register supports repeated data use, continuous updates, and serves as a basis for survey frames. (Eurostat 2021a)

Standards & authority services

IČO: The organisation identification number (IČO) is an identifier assigned to all types of legal entities, entrepreneurs and public authorities by the Statistical Office of the Slovak Republic. The Czech Republic's organisation identifier is also called IČO.

ISNI: an ISO certified global standard number for identifying the millions of contributors to creative works and those active in their distribution.

OpenCorporates: a public corporation database which sources data from national business registries.

VIAF: The Virtual International Authority File (VIAF) is an international service that consolidates multiple name authority files into a single database. Their primary goal is to enhance the efficiency and usability of library authority files by linking and merging widely used authority records and making them accessible online.

Statistical terms

The **Statistical Business Register** (SBR) is a foundational tool for statistical production (Eurostat 2021a)

universe: the full set of units that are relevant for a particular statistical study — the theoretical population from which observations could be drawn.

coverage: the subset of the universe that is actually included in the data collection (e.g., surveyed or recorded). It is used to assess completeness and representativeness of the data relative to the universe.

Data protection

GDPR: The General Data Protection Regulation (GDPR) is a legal framework made by the European Union that sets guidelines for the collection and processing of personal information from individuals who live in and outside of the European Union.

Data science terms

agent: a person, group, or organization capable of performing actions or making decisions within a system. In statistical and semantic models, agents are entities that generate, transform, or influence data. In our context, this includes natural persons (e.g., composers), corporate bodies (e.g., record labels), or informal collectives (e.g., bands). Adopted from (Wikimedia Deutschland 2025), in line with other event-based ontologies used in our work.

data: reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing Note 1 to entry: Data can be processed by humans or by automatic means. [SOURCE: ISO/IEC 2382:2015, 2121272] (ISO 2019a)

database: collection of data organized according to a conceptual structure describing the characteristics of these data and the relationships among their corresponding entities, supporting one or more application areas. [SOURCE: ISO/IEC 2382:2015, 2121413] (ISO 2019a)

data cube: A statistical data set created in a multi-dimensional space (e.g., time, geography, gender), or hyper-cube, indexed by those dimensions. The term cube shouldn't be taken literally, it is not meant to imply that there are exactly three dimensions.

data set or dataset: identifiable collection of data available for access or download in one or more formats [SOURCE: Adapted from ISO 19115-2:2009, 4.7] Beware: various conceptual and information models use different dataset definitions. (ISO 2019a)

interoperability: Ability of two or more systems or applications to exchange information and to mutually use the information that has been exchanged. [SOURCE: ISO/IEC 19941:2017] (ISO 2017a)

knowledge base: database that contains inference rules and information about human experience and expertise in a domain. 1: In self-improving systems, the knowledge base additionally contains information resulting from the solution of previously encountered problems. The terms knowledge base and K-base are standardized by ISO/IEC [ISO/IEC 2382-1:1993]. (ISO 2023)

knowledge representation: process or result of encoding and storing knowledge in a knowledge base. Term and definition standardized by ISO/IEC [ISO/IEC 2382-28:1995]. (ISO 2023)

knowledge graph: a knowledge representation that uses a graph-structured data model to represent and operate on data.

persistent identifier (PID): unique identifier that ensures permanent access for a digital object by providing access to it independently of its physical location or current ownership. [SOURCE: ISO 24619:2011, definition 3.2.4] (ISO 2017b)

structured data: data which are organized based on a pre-defined (applicable) set of rules
Note: The predefined set of rules governing the basis on which the data is structured needs to be clearly stated and made known. [(ISO 2019a)]

Organisations and abbreviations

SOZA: SOZA (Slovenský ochranný zväz autorský pre práva k hudobným dielam, Slovak Performing and Mechanical Rights Society) is a legal entity, non-profit civic association of authors and publishers of musical works, association of natural persons and legal entities.

Hudobné Centrum: Music Centre Slovakia is a music organisation with a mission to promote Slovak contemporaly music.

CISAC: the International Confederation of Societies of Authors and Composers is the world's leading network of authors' societies. It protects the rights and promotes the interests of creators worldwide.

Data name abbreviations

ISRC: International Standard Recording Code (ISO 2019b): a globally recognized identifier for individual sound recordings and music videos, used for tracking usage, reporting royalties, and rights management. It includes a country code, registrant code, year of reference, and designation code.

ISWC: International Standard Musical Work Code, a unique identifier assigned to a musical work. It allows rights management organizations to track the use and ownership of compositions across jurisdictions and platforms. Unlike ISRCs (which refer to recordings), ISWCs identify the underlying work regardless of performance.

ISNI: International Standard Name Identifier (Camp, Lieber, and IFLA 2022)

DOI: Digital Object Identifier, an international metadata standard for describing statistical and social science data. It supports the full data lifecycle — from conceptual design to data archiving — and includes models for documenting surveys, variables, code lists, and responses. (Data Documentation Initiative 2020b)

CIDOC-CRM: International Committee for Documentation–Conceptual Reference Model (Bekiari et al. 2024)

RiC: Records in Contexts–Conceptual Model (Archives Expert Group on Archival Description 2023)

DCTERMS or DCMI: Dublin Core Metadata Terms (Core 2020)

RDFS: Resource Description Framework Schema

EDM: Europeana Data Model (Europeana 2017)

ESCO: European Skills, Competences, Qualifications and Occupations (De Smedt 2023) is the multilingual classification system developed by the European Commission that identifies and categorises skills, competences, qualifications, and occupations across the EU. It complements ISCO and is available in both taxonomy and ontology form.

ISCO a hierarchical classification developed by the International Labour Organization (ILO) for organising jobs and occupations by type of work and skill level. It is used in labour statistics and aligned with ESCO at the EU level.

PROV-O: Provenance ontology (Lebo et al. "2013")

Other abbreviations

ECCCH: European Collaborative Cloud for Cultural Heritage. The ECCCH is a European Union initiative to create a shared digital infrastructure that connects cultural heritage institutions and professionals across the EU. It will provide specific digital collaboration tools for the sector while removing barriers for smaller and more remote institutions. (See the Commission Recommendation (EU) 2021/1970 of 10 November 2021 on a common European data space for cultural heritage (2021) for details.)

EOSC: The European Open Science Cloud is an environment for hosting and processing research data to support EU science.(Gilles 2023)

© 2023 OpenMusE | HORIZON-CL2-2022-HERITAGE-01-05 | Grant Agreement No. 101095295

Executive Summary

The European music ecosystem operates through fragmented, inconsistent, and often low-quality registration processes across rights management, cultural heritage, and statistical systems. This fragmentation increases administrative costs, reduces data reliability, and constrains the consistent measurement of the economic contribution of music. While large volumes of data are generated across these domains, they remain difficult to reconcile, validate, and reuse due to the lack of coordinated identifiers and interoperable structures.
D1.4 defines Open Music Registers as a coordinated, interoperable framework for identifying and maintaining information about music-related agents (authors, performers, producers, organisations) and their economic activities. The approach does not introduce centralisation or replace existing systems; instead, it establishes interoperability across existing registers, enabling shared use of data while preserving institutional autonomy and existing governance structures. In this sense, Open Music Registers function as a linkage and coordination layer rather than a unified database.
The deliverable presents:

- A conceptual model of agents, activities, and registers aligned with international statistical standards, including GSIM, SDMX, and DDI
- A data architecture for linking private and public data sources, including collective management organisations, business registers, and cultural heritage databases
- A definition of registers as the core infrastructure for economic measurement, including their role in statistical production, survey design, and indicator construction
- An interoperability-oriented approach aligned with European data space principles, enabling federated data use without requiring structural integration of source systems

The analysis demonstrates that current limitations in music sector statistics are primarily infrastructural rather than conceptual: relevant data exists across multiple domains but remains fragmented, inconsistently structured, and difficult to combine. By introducing shared identifiers, harmonised data models, and reproducible linkage mechanisms, Open Music Registers provide the conditions for improving data quality and comparability without imposing additional reporting burdens.
The implementation of Open Music Registers establishes:

1. Reduced duplication in registration processes through shared identifiers and coordinated data maintenance
2. Improved data consistency and interoperability across institutional domains
3. A stable foundation for producing comparable, timely, and methodologically consistent economic statistics

D1.4 provides the statistical and data-infrastructural foundation of the Open Music Observatory framework. Legal and policy considerations are addressed in D5.6, while D5.7 situates these components within a broader operational framework and elaborates their role in federated data infrastructures.

© 2023 OpenMusE | HORIZON-CL2-2022-HERITAGE-01-05 | Grant Agreement No. 101095295

Table of Contents

Introduction ... 11
1 Ambition: music registers that work together ... 13
1.1 Identification and Attribution Strategy in Open Music Registers ... 15
1.2 Preceding results from Open Music Europe WP1 ... 16
1.3 Harmonised registers for better research data collection and statistical production ... 17
1.4 Why a Dataspace? ... 18
1.5 Why Wikibase? ... 21
1.6 Forms of music or music entities ... 22
1.7 Technical harmonisation ... 23
2 Overview of the Currently Working Registers ... 25
2.1 Registering Data Subjects and Their Work ... 25
2.2 Music Publishers ... 25
2.3 KULT 5-01 ... 26
2.4 KULT 16-01 ... 28
2.5 KULT 19-01 ... 28
2.6 Composers ... 30
2.7 Performers ... 31
3 Surveying ... 32
3.1 Extending with GSIM ... 32
3.1.1 Identifying the subjects ... 33
3.1.1.1 Occupational and economic outcomes ... 33
3.1.1.2 Ex-post processing ... 34
3.2 Surveying the music economy ... 34
3.2.1 Open Music Europe WP1: Music Economy ... 34
3.2.2 Performing arts agents ... 34
3.2.3 Music publishing and sound recording agents ... 35
3.3 Surveying music diversity ... 36
3.3.1 Open Music Europe WP2: Music Diversity ... 36
3.3.2 Registration of musical entities ... 36
4 Conclusion ... 37
References ... 38
Appendix A — Open Music Economy Register ... 44
A.1 Value of the Data ... 44
A.2 Background ... 44

A.3 Collection statement 47
A.4 Data description 51
A.4.1 General concepts of surveying 51
A.4.2 General model of agency 53
A.4.3 Human agents: musicians, music professional 54
A.4.4 Informal corporate bodies: orchestras, informal enterprises 56
A.4.5 Musicians works and recordings 59
A.4.6 Formal corporate bodies: municipalities, enterprises, foundations 59
A.4.7 Business-to-business users of music 62
A.5 Experimental design, materials and methods 63
A.5.1 Occupations 64
A.5.2 Economic activities of corporations 69
A.6 Data Curation Policy for Slovak Music Economy Register 69
A.6.1 Data Collection and Ingestion 70
A.6.2 Limitations 71
A.6.3 Data Quality and Integrity 72
A.6.4 Data Ontologies, Taxonomies, and Enrichment 72
A.6.5 Data Updates and Maintenance 73
A.6.6 Data Access and Sharing 73
A.6.7 Data Security and Privacy 74
A.6.8 Governance and Responsibilities 74
A.6.9 Preservation and Archiving 75
A.6.10 Review and Continuous Improvement 75
A.7 Ethics statement 76
A.8 Acknowledgements 76
Appendix B — Appendix: How to use our registers 76
B.1 How to access our registers? 76
B.1.1 Human use 76
B.1.2 API Access 77
B.2 Who can access our registers? 77

Introduction

Registration is a necessary but burdensome process in the music sector. Musicians and their organisations routinely enter the same information—about people, works, recordings, or venues—into multiple systems: membership lists, royalty systems, event calendars, licensing databases, and more. The result is often duplication, outdated records, and growing cases of unregulated data scraping across platforms.

This report presents a way forward: by developing open, interoperable registers that can be shared across institutions, we reduce the effort of maintaining music-related metadata and increase its utility for statistics, rights management, and research.

ⓘ

This document has been prepared by OpenMusE project partners as an account of work carried out within the framework of the EC-GA contract no 101095295 (Open Music Europe 2023).

The Open Music Europe project has received funding from the European Union's Horizon Europe, research and innovation programme, under Grant Agreement No. 101095295*. 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.*

Any dissemination of results must indicate that it reflects only the author's view and that the Commission Agency is not responsible for any use that may be made of the information it contains.

This deliverable connects to a linked open data database supported by applications. As a living document conforming to OPA, it can be found at https://github.com/dataobservatory-eu/music-economy-register with standardised folders and file names.

📚 Live Documentation:

- HTML Book
- Download PDF or EPUB

A **register**, in IT terms, is a dataset that supports structured workflows. For example, a membership register should reflect new and departed members; a rights register should update when a composer dies and an heir inherits. Registers support specific applications: artists register with distributors to publish music, with CMOs to collect royalties, or with researchers to take part in surveys.

Maintaining these systems is time-consuming and costly. Many fail not for technical reasons, but because the expected benefits don't justify the effort required for complete and accurate registration. In some cases, conflicts of interest also discourage collaboration—such as when proprietary identifiers (e.g. ISRC, ISWC) are reused without oversight or cost-sharing.

The Open Music Registers project presents a new, practical approach to managing registration processes in the music sector. Today, musicians and organisations routinely submit the same data—about people, works, recordings, or venues—across many isolated systems. This redundancy increases administrative costs, creates inconsistencies, and leads to outdated or fragmented records.

Open Music Registers proposes a solution: a **shared, open, and interoperable infrastructure** where carefully harmonised registers can serve multiple purposes at once—rights management, statistics, cultural policy, research, and business innovation.

Building on the foundations of the *Open Music Europe* project [Open Music Europe (2023), we apply modern data governance standards (such as the European Interoperability Framework, FAIR principles, and dataspace design patterns) to streamline registration processes while respecting GDPR and domain-specific requirements. We rely on technologies like **Wikibase**—the platform behind **Wikidata**—to create flexible, multilingual, standards-compliant knowledge bases.

Our work focuses initially on Slovakia, where we have constructed a live, updateable **Slovak Music Economy Register** that covers:

- Composers, publishers, performers, and producers,
- In an exploratory phase, venues and events,
- Business entities linked to cultural and economic activities.

In the WP2 we are creating a comprehensive database of sound recordings and musical works connected to the composers, performers, publishers, and record labels.

We align these datasets with international standards (ISNI, ISRC, ISWC, OpenCorporates, ESCO, NACE, ISCO) to ensure interoperability with statistical offices, rights organisations, libraries, and platforms across Europe.

Our system is designed not as a centralised database, but as a **federated dataspace** that respects the autonomy and ownership of each contributor while enabling smarter data reuse, discoverability, and analysis. It integrates seamlessly with European initiatives like the European Collaborative Cloud for Cultural Heritage (ECCCH) and the European Open Science Cloud (EOSC).

The ultimate goal is to reduce the burden on artists, improve the quality and visibility of music metadata, and enable cultural and economic indicators that are currently missing from national and EU statistics. By building a multi-purpose, standards-aligned infrastructure, we help bridge the worlds of cultural heritage, the creative industries, and official statistics—laying the groundwork for more inclusive, efficient, and evidence-based policymaking in the European music sector.

This documentation is structured as follows:

- We introduce the conceptual and policy background for Open Music Registers (Chapter 1);
- We present how can the user use the existing registers (Chapter 2);
- And we outline how it can be used to support administrative and sample surveys in cultural statistics (Chapter 3);
- A detailed technical annex shows the data model and modeling considerations (Appendix A);
- Another Annex gives further information on how to use our system, and provides some overview of the register used for WP2 about the diversity and circulation of musical works and their print and recorded manifestations.

We touch only briefly on the issue of surveying musical diversity, which will be addressed in a separate report. That work focuses not only on agents—composers, producers, performers—but also on their creative outputs: works, lyrics, recordings, and editions.

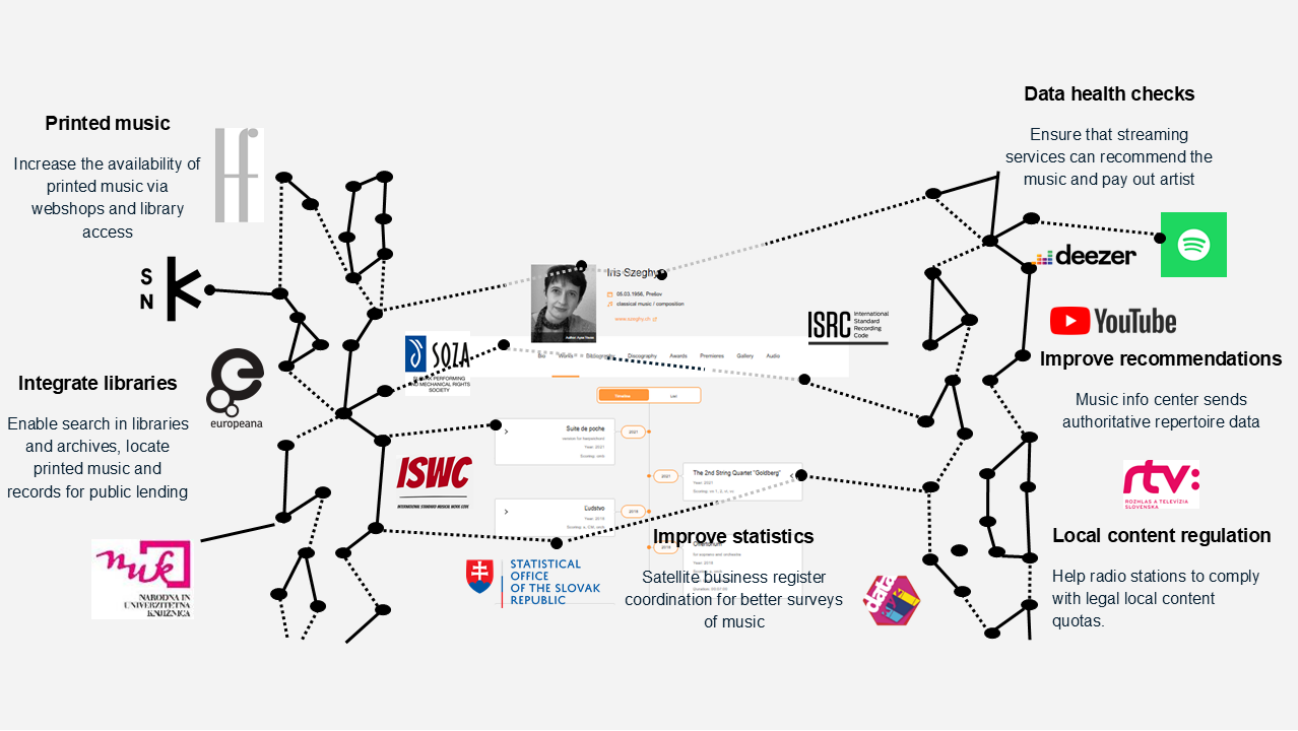
1 Ambition: music registers that work together

When someone registers a new person or entity in a system, they are identifying a subject and attaching key facts about them to a consistent record. Today, such records are almost always managed through spreadsheet tables or relational databases. However, as digital services multiply, maintaining parallel registries for every new purpose has become a tedious, error-prone, and increasingly problematic practice. Artists, their groups, and their associations are asked to submit more and more structured data — often without full oversight or control over where and how this information is stored or reused.

Our vision is a **shared registration process**: a unified platform where an artist can fill out a single form to submit the necessary data and documents, grant GDPR-compliant consent, and manage their information over time. Through this interface, they could register and maintain data for multiple purposes — including rights management, public funding, and broader visibility through open knowledge platforms.

This shared process would support data needs across: - collecting societies - cultural ministries - national statistical institutes - international databases like Wikidata or MusicBrainz - streaming and discovery services

The artist remains the source of truth. With proper consent and traceability, each registration contributes to a connected ecosystem.

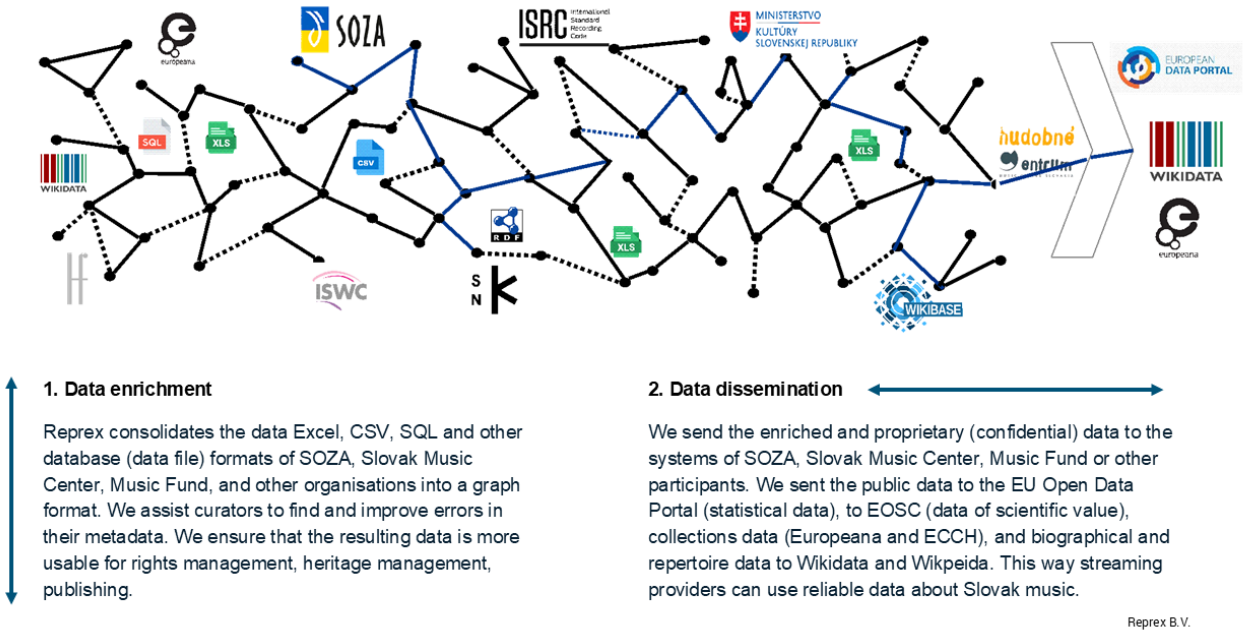


The graph represents connections between data nodes (e.g. database tables or fields) mediated by the Open Music Observatory. Solid lines indicate bidirectional or mutually maintained data exchange between systems (e.g. where both sides can update and synchronise information). Dotted lines indicate unidirectional or restricted data flows (e.g. platform-specific access such as streaming data provided by services like Spotify or Deezer without reciprocal data integration).

The composer Iris Szeghy was one of the first persons who allowed her data to be used in our shared register. With her example we show how the coordination or registration processes can lead to better music services for music users and researchers. See our presentation for the International Association of Music Information Centers.

This illustrates how a harmonised registration process can support both cultural recognition and reuse.

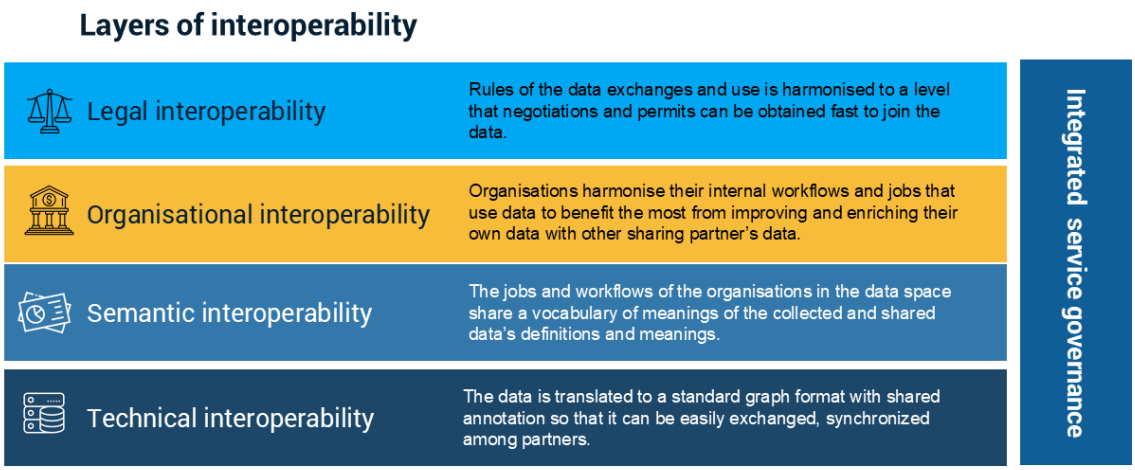
If you are a self-releasing artist or artist manager, you probably fill out dozens of registration forms each year — for databases, web platforms, grant applications, copyright management, and more. Over time, it's easy to lose track. And while you may forget these forms, your data doesn't: it lives on, often duplicated, outdated, or disconnected — scattered across the internet. It's a bit like space debris: years of untracked, obsolete data floating around. Each new registration risks colliding with the old unless registers are connected and responsibly maintained.



Our systems allows a vertical data exchange, which allow that trusted, cooperating music organisations cross-check their data about artists and their works, recordings, releases, and then make them available for streaming providers, festival promoters, radio editors and other users.

The time has come for technology that connects these registers, returns control to the artist, and enables a single, authoritative point of truth for reliable, maintained data.

This system should not only save time for both artists and service providers, but also increase transparency, improve data quality, and support informed consent in the management of professional and personal information. In the longer term, it could enable regular reviews, help artists monitor where and how their data is used, and offer feedback on the quality and discoverability of their public profile.



We use the European Interoperability Framework as a general conceptual guideline to create a multi-purpose, interoperable registration process.

This ambition builds on earlier research and practical experience, which we summarise in the next section.

1.1 Identification and Attribution Strategy in Open Music Registers

The Open Music Registers initiative is grounded in a pragmatic approach to managing the balance between identification and attribution of entities within the music economy. We recognize that both processes are essential but costly and potentially error-prone. Therefore, our strategy is to minimize duplication, maximize reuse, and support reliable identification processes wherever feasible.

Identification refers to recognizing and linking entities—such as musical works, rightsholders, and recordings—to existing, trustworthy identifiers. Where reliable identification services are in place (e.g., ISNI for persons, VIAF for authority control, ISWC for musical works managed locally by SOZA), we prioritize integration and reuse. We aim to support identification efforts, for example, by assisting SOZA in reducing the costs of work identification and helping the Slovak National Library verify manifestations through structured metadata contributions.

Attribution involves actively assigning identifiers, typically when no existing trusted identifier is available. When necessary, the Open Music Registers facilitate attribution processes, but always with care to follow international best practices and to avoid unnecessary parallel identification efforts.

In practice, metadata management costs in the music sector are often disproportionately high relative to the economic value of the assets. Many musical works generate only a few cents in royalties per month—or even less—while the costs of accurate identification, attribution, and metadata maintenance can easily exceed these earnings. The music ecosystem is exceptionally complex: hundreds of millions of assets are connected to millions of rightsholders, whose circumstances (names, publishers, labels, heirs) frequently change.

This complexity contributes to systemic problems in royalty distribution, rights management, and cultural visibility. Without cost-effective metadata solutions, the economic viability of accurate rights

management becomes unsustainable, especially for small rightsholders and emerging artists. Left unchecked, this creates a vicious circle: assets remain under-identified, royalties are misdirected or delayed, trust in the system erodes, and the overall cultural economy suffers.

The Open Music Registers project is designed to break this cycle by:

- Trusting and reusing validated attribution results from authoritative registrars.
- Undertaking new identification only when necessary and justified.
- Avoiding duplication and parallel registration processes.
- Lowering the burden on artists and registrars alike through interoperability and shared infrastructure.

By strategically balancing identification and attribution, and by building efficient, scalable metadata solutions, we aim to ensure that music assets can be properly valued, fairly remunerated, and sustainably managed—even when the economic value per asset is modest.

1.2 Preceding results from Open Music Europe WP1

This work builds on the statistical and methodological foundation laid by the Open Music Europe project. One of the project's central insights is that, although digital metadata about the music economy is increasingly abundant, core economic indicators — such as employment, income, and production investment — are still poorly captured in national and European statistics.

Two former deliverables helped map this gap. The first, a European-wide methodological assessment D1.1 (Antal, Kmety Barteková, and Remeňová 2023), presented in Bratislava in 2023, showed that fragmented data collection practices limit the use of standard economic codes such as NACE and ISIC. It recommended building open, artist-facing infrastructures that allow music professionals to register their data once and have it reused across multiple systems — for research, rights management, or policy analysis.

The second study (Antal 2023), demonstrated how data could flow between music organisations, statistical offices, and national registers. It reused earlier national surveys — KULT05, KULT16, and KULT19 — which had gathered useful information about music creators and organisations. However, those surveys lacked persistent identifiers and mechanisms for updates over time.

The Slovak Music Economy Register — developed under Task T1.2 — responds directly to this challenge. It aims to maintain a live, updateable register of music professionals and organisations, combining survey data, rights management records, and self-registered entries. It is part of the larger Slovak Comprehensive Music Database (SKCMDb), which also contains registers of musical works, sound recordings, and performance venues. Together, these connected datasets support better cultural statistics, rights attribution, and policy design.

By integrating sources like SOZA (the Slovak collective management organisation) and the Slovak Music Centre, the register creates a flexible but standards-compliant way to track economic activity in the music sector — one that respects national statistical needs while also supporting transparency, interoperability, and creator rights.

1.3 Harmonised registers for better research data collection and statistical production

In statistical terms, a *register* is a structured list or database that systematically tracks units of interest—typically enterprises, individuals, or institutions—over time. Public statistical authorities often maintain *population registers* or *enterprise registers* that serve as the backbone for sampling, indicator production, and policy monitoring.

However, in many industries, particularly cultural and creative ones, official registers are incomplete or too broad, grouping diverse sectors under umbrella codes like NACE J58 or R90. This is partly because the typical organisation size in these industries is smaller than the enterprise size that participates in traditional, mandatory sample-survey statistical production, and partly due to the high level of informality in these sectors, as reported in T1.1 see (Antal, Kmety Barteková, and Remeňová 2023; Antal 2023). These classifications dilute the specificity needed for meaningful data analysis. For instance, a music enterprise might be lumped together with film postproduction, making the resulting statistics less actionable for business decision-making or targeted public policy.

In contrast, private organisations often maintain richer, more detailed internal registers for their operations. These might include:

- *Membership lists* maintained by professional associations or unions, such as those representing music publishing entities or record labels. For example, we created a Slovak register of music publishers; under the collective management directive (European Parliament and European Council 2014), such registers must be made public, even if not well-structured.
- *Client databases* used by CMOs or licensing agencies, detailing entities like restaurants that play music. While these databases are proprietary, public business registers or local registers of licensed venues (e.g., restaurants, music clubs requiring local permission due to noise and traffic considerations) can help create connected music user registers.
- *CRMs* (Customer Relationship Management systems) used by businesses for managing partners and contractors.

While these private registers are typically well-aligned with operational needs, they are not harmonised with national statistical systems—meaning identifiers (e.g., tax IDs, sector codes, geographic location data) may not match, formats vary, and there's no common protocol for data structure or update frequency.

Harmonisation, in this context, does not mean full integration but rather the voluntary alignment of **key reference variables** and **metadata**:

- Using public identifiers (e.g., business registry numbers),
- Mapping internal categories to standard classifications like NACE or ISIC,
- Agreeing on common reference periods and geographical codes.

With minimal effort, private actors can enhance the *statistical usability* of their data while still serving their internal purposes. This opens the door for *interoperability*, *cross-validation*, and more efficient data sharing—all without relinquishing control over proprietary data. We will illustrate these efforts in the next section of this document.

Efforts to harmonise registers across private and public actors also create the foundation for more advanced and resilient forms of **statistical production**. As previously noted, when data collection and processing can happen closer to the data source—whether through sampling, microdata preparation, or anonymised outputs—we reduce both privacy risks and operational burdens. This distributed model of statistical production reflects an important shift: the production of public value is no longer confined to the walls of official statistical offices but extended into the data infrastructures of cooperating institutions.

These changes are supported and legitimised by a growing ecosystem of **European data governance frameworks**, including the *Data Governance Ac* (European Parliament and Council of the European Union 2022) and *Data Act* (European Parliament and Council of the European Union 2023), which encourage secure and fair reuse of privately held data for public interest purposes. More ambitiously, the development of **European Common Data Spaces**—such as the *Cultural Heritage Data Space* and upcoming *Media and Creative Industries Data Space*—aims to build sector-specific environments where data can be shared, pooled, and reused under common rules and interoperable standards. In *Open Music Europe*, we aim at the creation of a small, extendible music data space in the form of the *Open Music Observatory* that can federate data with other parts of the European data spaces.

Together, these legal, technical, and institutional innovations represent a new chapter in the co-production of knowledge: one where governments, cultural institutions, and private sector actors jointly curate the evidence base needed to shape smarter, more responsive policy.

1.4 Why a Dataspace?

In building the Open Music Registers, we deliberately avoided merging different data sources into a single centralised database. This decision reflects the lessons learned and best practices laid out in the *UN Guidelines on Statistical Business Registers* (United Nations Statistics Division 2024), and the microdata coordination of the Dutch statistical authority, CBS (Sun 2021).

A register aims to be a complete list of the objects in a specific group or population (Anders and Britt 2007). Norway created statistical registers to tap into governmental data stores in 1990 and into municipal ones in 1995; by 2019, it utilised about 100 records and drew data from 30 public institutional sources (Sæbø and Dimakos 2023, 1). Like most countries, such “administrative data” was retrieved from other governmental entities, not the private sector, and at the same time replaced almost all sample surveys which had inferior data qualities.

A statistical business register (SBR) is a regularly updated, structured database of economic units in a territorial area, maintained by a national statistical office, and used for statistical purposes.

In official statistics, the Statistical Business Register (SBR) is the foundational frame for producing core statistical products: enterprise statistics, labour market statistics, and ultimately, derived indicators like the GDP. As the Guidelines emphasise:

“High-quality business statistics depend on a high-quality SBR. A high-quality SBR fulfils the user needs in an optimal way, and is based on international concepts, definitions and classifications. Thus, it serves as the basis for international harmonisation of economic statistics in terms of coverage, statistical units and frame methodology.” (United Nations Department of Economic and Social Affairs 2024, p2)

Building a statistical business register is not a mechanical task of merging administrative sources.

"Even if the SBR is based on information from administrative registers, the concepts, characteristics, methods of maintenance, and other aspects of the SBR need to be based on statistical concepts." (United Nations Department of Economic and Social Affairs 2024, p4)

Official statisticians must carefully construct statistical units from diverse administrative sources — such as tax records, company registries, and social security data — which differ in legal basis, workflows, update cycles, and semantic structures.

Similarly, in Open Music Registers, our sources include copyright registration systems, library authority files, repertoire asset lists, and royalty accounting databases. These sources have different legal status, different operational goals, and different metadata models. Therefore, just as statistical agencies cannot simply merge administrative registers, we cannot simply merge these music-sector registers.

Instead, our approach mirrors that of official statisticians: harmonisation through shared semantics and carefully coordinated data linkage.

In official SBRs, the statistical office typically has legal authority to impose common taxonomies, data standards, and reporting obligations. This greatly simplifies harmonisation, but even then, good cooperation is required by statistical authorities, too.

"Technical and organisational issues hampering access by national statistical offices to administrative databases can be addressed by establishing good cooperation with administrative bodies, ideally formalised in memoranda of understanding." (United Nations Department of Economic and Social Affairs 2024, p6)

In contrast, a data sharing space like Open Music Registers cannot rely on hierarchical authority. Instead, we operate under the voluntary, consensus-based model typical of data spaces:

- Metadata semantics (the definitions of terms and structures) are pre-agreed among participants.
- Actual data exchange happens only on an as-needed, as-permitted basis, respecting data licences, contractual rights, and privacy regulations.
- Differences in workflows and purposes — for example between a rights management organisation and a national library — must be reconciled collaboratively through decentralised negotiation, not imposed centrally.

This makes data harmonisation slower, but it also ensures higher resilience, broader participation, and stronger trust among stakeholders — goals very much aligned with the emerging standards for European data spaces and FAIR data principles.

Dataspaces: Fundamentals, Principles, and Technique: "A dataspace is an emerging approach to data management which recognises that in large-scale integration scenarios, involving thousands of data sources, it is difficult and expensive to obtain an upfront unifying schema across all sources. Data is integrated on an `as-needed' basis with the labour-intensive aspects of data integration postponed until they are required. Dataspaces reduce the initial effort required to set up data integration by relying on automatic matching and mapping generation techniques." (Curry 2020)

Dataspace for Cultural and Creative Industries (Position paper): "A data space is an ecosystem of exchange, processing, sharing and provision of data between trusted partners, for a fee or not. It is not about copying or repatriating data centrally but about ensuring that each data holder has full

control over the conditions (e.g., who, when, and under what condition) of access to their data." (EBU and Gaia-X 2022, p16)

As we developed the Open Music Registers, we recognised that traditional database architectures would be insufficient for our goals. Our challenge resembled those faced by national statistical offices: the metadata about music agents, works, and performances is dispersed across many organisations, collected under different legal mandates, maintained with diverse standards, and updated for different operational needs. Following emerging best practices in official statistics, particularly those demonstrated by Statistics Netherlands (CBS), we chose to create a knowledge graph-based infrastructure (Sun 2021); by doing so, the CBS has de facto implemented a data sharing space structure.

Instead of imposing centralised control over the sources, we harmonise their metadata using shared ontologies (such as DCTERMS, CIDOC-CRM, and music-specific standards like ISWC, ISNI, ISRC) and link the underlying datasets through interoperable, semantic structures. This approach enables flexible, incremental integration across the music sector, while preserving the autonomy and legal rights of each data contributor.

The CBS Microdata project offers a clear parallel to our ambition. Faced with fragmented, departmentally managed microdata collections, CBS built a knowledge graph to unify metadata across its research datasets, using standard vocabularies such as DCAT and Dublin Core Terms. They extracted key information from separate PDF documents, mapped variables to standard ontologies, and created an RDF-based knowledge graph to support researchers seeking to combine and explore disparate datasets. Like CBS, we aim not merely to connect databases, but to create an integrated semantic environment where different registration workflows—rights management, public libraries, music asset lists—can be understood and cross-referenced. Just as CBS facilitates high-quality, FAIR-compliant research across domains like health, education, and employment, Open Music Registers will support rights management, cultural research, policy evaluation, and statistical production across the music economy.

Our ambition does not stop here. Statistical authorities increasingly realise that cooperation with data-rich private sector organisations can improve the quality of official statistics. Why ask imprecise and boring consumer purchase diaries about household holding when credit-card companies have exact data about more than 95% of all purchases? Statistics Norway realised that instead of asking 7,000 households what they were buying in the supermarket, it is far more accurate and potentially cheaper to acquire the data directly from the sales logs of the supermarkets or the payment transaction records of the credit card companies. Ask indirectly household energy use when the electricity trading and distribution system has the energy use data of every building and household? French statisticians prefer this data source.

As explained in our preceding T1.1 (Antal 2023), the data space aproach, and the utilisation of the same metadata standards, allow a potential cooperation with Slovakia's cultural satellite accounts. For this, we needed to follow recent international guidelines on using privately held data for official statistics (UNECE 2015; United Nations Statistics Division 2020; Eurostat 2021b), and more broadly from the policy and research literature on business-to-government data sharing (ESS 2017; "Towards a European Strategy on Business-to-Government Data Sharing for the Public Interest. Final Report Prepared by the High-Level Expert Group on Business-to-Government Data Sharing" 2020; Vigorito 2022; "Business-to-Government Data Sharing for Public Interests in the European Union: Results of a Public Consultation" 2022). Such complex legal, security and organisational requirements also help design similarly high-quality processes when sharing data among private partners.

1.5 Why Wikibase?

Wikibase is the open-source software platform that powers Wikidata — the largest open knowledge graph on the internet. It enables organizations to create their own structured, linked, multilingual databases, while remaining interoperable with global ecosystems like Wikidata. By using Wikibase, Open Music Registers can build an independent but connected knowledge base that integrates identifiers, metadata, and factual information from many sources, while retaining control over the quality, governance, and authority of the data.

We are building a **knowledge base** that enables the combination of data, information, and knowledge about music in novel and valuable ways. This requires us to integrate resources that were traditionally managed separately, each within different professional, commercial, or cultural workflows. As a result, we must work with identifiers developed by various stakeholders concerned with the production, commercialisation, documentation, and preservation of manifestations of music, or knowledge about music — whether it concerns printed scores, MIDI files, sound recordings, audiovisual content, or books about music.

Every domain — whether music, publishing, research, business, or cultural heritage — has developed its own identification systems, each optimised for the specific needs of that sector. Following the evolution outlined by Laura Manzoni, the development of modern identifiers reflects a growing need for persistent, stable, and interoperable ways to *identify resources*, not only online, but across the entire information infrastructure (Manzoni 2022).

- **ISBN (International Standard Book Number)** was introduced in 1967–1970 to uniquely identify books and other monographic publications. It ensures that each edition of a book can be cited and tracked across publishers, libraries, and sales platform.
- **ISMN (International Standard Music Number)** followed to identify printed music (sheet music scores) across publishers and distributors. It originally started a numeric range of ISBN, but by now printed sheet music has its own registration system (ISO 2021). We work with print music diversity and circulation in WP2.
- **ISRC (International Standard Recording Code)** was created in 1986 to uniquely identify individual sound recordings and music video recordings, independent of their format (CD, vinyl, digital download). It tracks usage and royalties across rights management systems (ISO 2019b; Camp, Lieber, and IFLA 2022). We work with recorded music diversity and circulation in WP2.
- **ISWC (International Standard Musical Work Code)**, introduced in the 1990s, identifies musical works (i.e., the underlying composition or lyrics) rather than any particular recording or performance. It is used by collective management organisations (CMOs) to track royalties and rights across performances, broadcasts, and reproductions (ISO 2022).
- **DOI (Digital Object Identifier)**, established in the late 1990s, providing persistent identification of digital objects (articles, datasets, reports, etc.) independently from where they are hosted. Unlike a URL, the DOI remains stable even if the resource moves.
- **ISNI** identifies individual creators, performers, and contributors.
- **ISBN** identifies books and printed editions (including music books) .
- **OpenCorporatesID** identifies legal corporate bodies and music business organisations.

© 2023 OpenMusE | HORIZON-CL2-2022-HERITAGE-01-05 | Grant Agreement No. 101095295

These different identifiers capture different aspects of musical and cultural reality. They serve different workflows — such as royalty collection, publishing, research citation, market transparency — and operate under different legal and technical frameworks. If we want to *increase the interoperability* of services and systems that use music metadata — including rights management, cultural heritage, statistical analysis, or AI-driven discovery — we must ensure that *the identification of the same thing or person across contexts is connected*.

This is why we are not building a "single system" of identification, but rather a hub of connections between identifiers, anchored by careful mappings between authoritative registers, and supported by persistent, machine-readable links.

In this model, Wikidata functions as an **identity broker** — a global open knowledge graph that aggregates and connects identifiers from many domains, linking them to structured factual data.

By following this approach, Open Music Registers does not replace existing authority systems, but connects them intelligently, allowing music agents, works, performances, recordings, and organisations to be discovered, verified, linked, and reused across multiple platforms and sectors.

In recent years, Wikidata's role as an identity broker has been recognised and increasingly trusted by major institutions in cultural heritage, education, and research.

Studies have shown that UNESCO, Creative Commons, national libraries, research funders, and large-scale digital humanities projects actively integrate Wikidata identifiers to improve metadata quality, enable multilingual access, and facilitate cross-domain interoperability (van Veen 2019; Bianchini, Bargioni, and Pellizzari di San Girolamo 2021; Bianchini and Sardo 2022; Fagerving 2023). This growing trust stems from Wikidata's capacity to link persistent identifiers from different authoritative systems — such as VIAF, ISNI, DOI, ISBN, and institutional databases — into a public, machine-readable knowledge graph that supports both human discovery and automated services at a global scale.

In fact, it shows that there is no need for one standardization of practices for establishing the headings and structure of authority records in one international form; instead, users' convenience can be achieved by a technological infrastructure capable to present to each user the information about an entity in its own language and script. Additionally, Wikidata is the most evident example of the distributed and diffused approach of the semantic web to the issue of the universal identification of the entities. Also, Wikidata identification and description show that authority and bibliograp

This approach is now also reflected in major European policy initiatives. The EU Knowledge Graph, developed as part of the European Commission's linked data strategy, is building a cross-domain graph that explicitly integrates references from Wikidata, authority files, open registries, and official databases, recognising the need for semantic mediation across scientific, cultural, economic, and administrative domains (Diefenbach, Wilde, and Alipio 2021). Rather than replacing local or domain-specific authority files, Wikidata — and projects modelled on it — connect them, acting as a semantic bridge across fragmented information ecosystems.

This strategy is particularly powerful for cultural and creative domains like music, where different aspects of the same reality (creative works, recordings, performances, corporate actors) are registered in specialised, non-overlapping systems.

1.6 Forms of music or music entities

Our primary interest is not necessarily in the data subject but in their music, particularly their work (the occupation, artistic or economic activity they are engaged in) and its result in new music entities, such as works and their recorded and print manifestations. Copyrights and neighbouring rights connected to these entities, and their revenues. Event where and when these works are recorded or performed for an audience.

We are particularly interested in a parallel registration of works and the recorded fixations of their studio and staged performances intended for the public. (For example, we do not want to provide interoperability for recording information about rehearsals or liturgical uses of music for spiritual or faith-based communities.)

Subject	What is their role?
musical work	Composed of a combination of sounds, with or without accompanying text (ISO 2022)
performance	
live performance	The performance of music in front of a live audience.
recorded performance	The performance of music, interpreted from a musical work or improvised, with the aim of creating a recording.
sound recording	A recorded fixation, released for commercial or non-commercial use, an re-playable object from a recorded performance.
music venue	A building, a publicly accessible site, a piece of infrastructure or built structure where music is played for the public.

1.7 Technical harmonisation

Private entities can enhance the statistical usability of their data with relatively small effort, maintaining their internal operational needs while enabling interoperability, cross-validation, and efficient data sharing. This approach allows for collaboration without compromising control over proprietary information. The Open Music Observatory and the Slovak Comprehensive Music Database (SKCMDb) exemplify innovative methods of semantic data integration, bridging statistics and private cultural data infrastructures.

We employ Wikibase, an open-source platform underpinning Wikidata, to create a “private Wikidata” environment. This setup offers high visibility, reusability, and collaborative potential. By leveraging a widely adopted, multilingual knowledge graph, we lower barriers to integration and reuse, benefiting statistical agencies and cultural researchers alike.

Mapping data from organizations such as SOZA, the Slovak Music Centre, the Statistical Office of the Slovak Republic, and the Slovak National Library to standard classifications like ESCO, SDMX, NACE, and ISCO establishes semantic connections between internal datasets and international frameworks. This strategy enhances statistical interoperability, supports multilingual capabilities, and facilitates automated updates, making the data more accessible and AI-ready.

The Open Music Observatory not only links music industry business data but also integrates cultural, labour, and economic dimensions. This comprehensive approach aligns with the objectives of cultural

satellite accounts and Common Data Spaces initiatives. In this deliverable, we demonstrate how measuring the labour and cultural aspects of music can be supported through data infrastructures based on shared registration processes and harmonized registers.

2 Overview of the Currently Working Registers

The Open Music Registers initiative is not just a theoretical proposal. We have already implemented key components of a working system, notably the Slovak Music Economy Register and its broader context within the *Slovak Comprehensive Music Database* (SKCMDb).

The Slovak Music Economy Register is designed as a live, updateable register of music agents active in the Slovak Republic. It currently covers more than 3,000 agents, including natural persons (such as composers, performers, producers) and corporate bodies (such as music publishers, record labels, music groups, and venues). Each agent is connected, where possible, to persistent identifiers (such as ISNI, VIAF, or OpenCorporates IDs) to ensure interoperability and traceability across systems.

These entities are captured through an event-based, multilingual, semantically harmonised knowledge graph structure, with technical foundations in Wikibase (the same platform underlying Wikidata). This approach allows not only cross-validation of records, but also dynamic, machine-readable connections between creators, their works, the recordings of those works, and the venues where they are performed.

2.1 Registering Data Subjects and Their Work

One of the core functions of any **register** — whether it's a business register, a music economy register, a library authority file, or a rights management database — is to identify entities (people, works, companies, recordings, events, etc.) in a stable, unique, and reusable way.

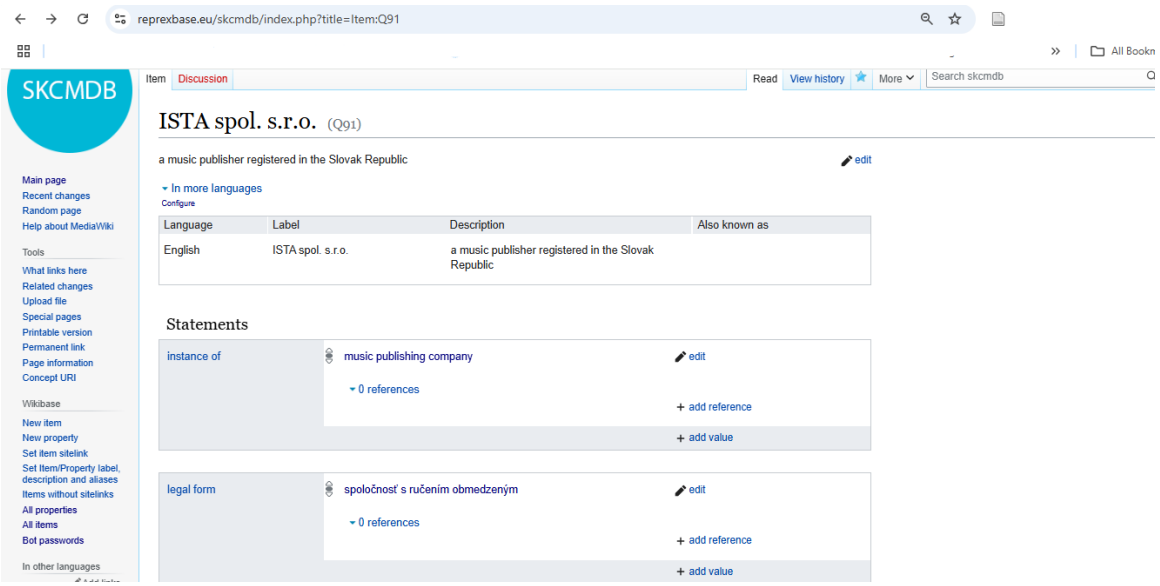
Modern registers are based around permanent identifiers, which are designed to be stable, unique, long-lasting — it identifies a thing even if the location changes. It is often just a unique string.

- ISNI identifies individual creators, performers, and contributors (Camp, Lieber, and IFLA 2022).
- OpenCorporatesID was introduced as part of the OpenCorporates project (founded 2010) to uniquely identify companies and corporate entities worldwide. It assigns a machine-readable URI to each registered company, enabling better transparency, traceability, and interoperability of business data.

Manzoni notes that the shift from identifying physical objects to identifying intellectual and legal entities — and later to assigning persistent identifiers to digital metadata and real-world organisations — reflects the expansion of the semantic web and the linked data ecosystem. Today, good data management practice across cultural, academic, and business domains depends on using *PIDs expressed as URIs* to ensure *long-term discoverability, interoperability, and trust*.

In the 2020s, modern permanent identifiers usually take the form of an URI, which is a technical format that is based on the identification infrastructure of the world wide web. An URI is a more general form of an URL; a web browser or a web application can work with it. Unlike an URL, that is a unique identifier for a digital document, an URI can identify a physical object, a living or deceased person; like the Library of Congress URI https://id.loc.gov/authorities/names/n79070061 for "Igor Stravinsky".

2.2 Music Publishers



The entry of ISTA spol. s.r.o. includes its registration number in the official Slovak business register and in OpenCorporates

You can query the register of all orchestras and ensembles:

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX wd: <https://reprexbase.eu/entity/>
PREFIX wdt: <https://reprexbase.eu/prop/direct/>
SELECT ?item ?itemLabel
WHERE {
 # Find the class entity by label (e.g., "music group")
 ?class rdfs:label "music publishing organisation"@en .
 # Find all items that are instance of that class
 ?item wdt:P5 ?class . # P2 = instance of (adjust this if different in your setup)
 OPTIONAL { ?item rdfs:label ?itemLabel . FILTER (lang(?itemLabel) = "en") }
}
LIMIT 100
```

Try it out⇗

Here will be a working link to our database with a screenshot and filter only ensembles that are part of the KULT 5-01 register.

And filter only ensembles that are part of the KULT 5-01 register.

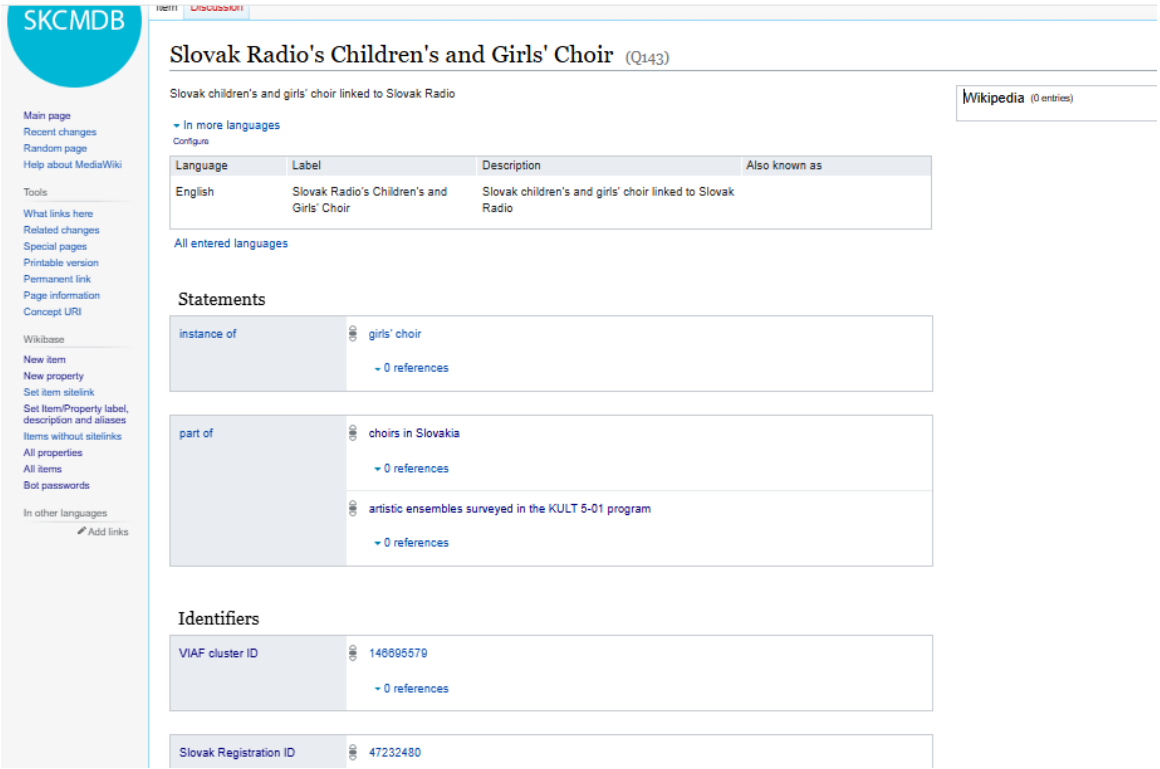
2.3 KULT 5-01

The KULT-5-01 survey is an official statistical instrument mandated by the Ministry of Culture of the Slovak Republic and registered by the Slovak Statistical Office (Ministerstvo kultúry Slovenskej republiky 2022c). It is conducted annually under the legal framework of the State Statistics Act No. 540/2001 Coll (Národná rada Slovenskej republiky 2001). The survey collects detailed data on musical ensembles and performing arts groups, focusing on their activities, audiences, financial performance, and human resources. Typical questions cover the number and genre of concerts performed (in Slovakia and abroad), repertoire composition (works by Slovak and foreign authors), visitor statistics (including performances for children and youth), economic data on income and

expenditure (broken down by grants, sponsorships, ticket sales, and other sources), and workforce composition by employment status, education level, and gender.

In previous work — notably the *Feasibility Study on Promoting Slovak Music in Slovakia & Abroad*, on which WP1 and WP2 are partly based (Antal 2020), and the subsequent D1.1 deliverable of WP1 *European-wide methodological assessment* (Antal, Kmety Barteková, and Remeňová 2023) — we highlighted the extreme difficulty of surveying the *informal economy* in cultural sectors. Many creators operate without formal business structures, and traditional sampling frames (like business registers) exclude them entirely.

Most music is performed by a group of performers, a musical group, which can take many forms: a duo, a trio, a music ensemble, orchestra, a rock band. Sometimes membership in these corporate bodies is long-standing, sometimes members change frequently. Musical groups and orchestras of professional musicians perform both economic and artistic activities; however, it is rare that the members of the orchestra also from a completely overlapping, membership based enterprise form, too. We say that muisical groups are often part of the informal economy because their economic performance often cannot be connected to a well-defined business entity.



The Slovak Radio's Children's and Girls' Choir does have a Slovak IČO identifier, but that is the identifier in the business register of the entire public broadcaster. The VIAF ID (managed by the Slovak National Library as control authority) is a far more suitable identifier for this choir. In this case, we also have to mention that the choir is made of minors, further complicating linking the artistic performance to the results of its economic dimension.

Through the Open Music Registers, we have registered all musical ensembles that submit KULT 5-01 forms, allowing researchers to easily identify which groups are officially surveyed and which are not. In cooperation with the Ministry of Culture, we have also received several years of historical microdata from the KULT 5-01 survey. This enables the creation of new statistical collections where, by clearly filtering ensembles already registered through KULT 5-01 and using our broader,

harmonised register, researchers can construct retrospectively consistent indicators for audience development, performance measurement, and cultural policy analysis. It also makes it possible to generate improved, partly backcasted statistics on a larger, more representative survey frame.

Two related services, connected to our Open Music Registers will ensure that more and more musical groups are properly identified. With our cooperation with the Slovak National Library, we are informing them about entities with curated metadata fields that should have a VIAF ID. Our UNLABEL service, developed in Open Music Europe by ALOADED and REPREX aims to give better data representation to self-releasing artists and provide them with an ISNI identifier.

Additionally, based on our harmonised metadata model, we are able to provide targeted recommendations for improving future editions of the KULT 5-01 survey, such as better classification of informal and semi-professional groups, capturing underreported repertoire, and collecting more detailed and structured data on financial flows and employment patterns.

2.4 KULT 16-01

he KULT 16-01 survey is an official statistical instrument mandated by the Ministry of Culture of the Slovak Republic and registered by the Slovak Statistical Office (Ministerstvo kultúry Slovenskej republiky 2022a). It is conducted annually under the legal framework of the State Statistics Act No. 540/2001 Coll. The survey collects detailed data on public music events organised professionally across Slovakia. It focuses on concert activities, audience attendance, event types, financial performance, and human resources. Typical questions cover the number of concerts by genre (classical, jazz, rock, folk, etc.), their regional distribution, visitor statistics (including paid and unpaid entry), and the organisation of periodical and non-periodical music events such as festivals, concert series, and competitions. It also collects detailed financial data on revenues and costs associated with event organisation and information on the human resources involved.

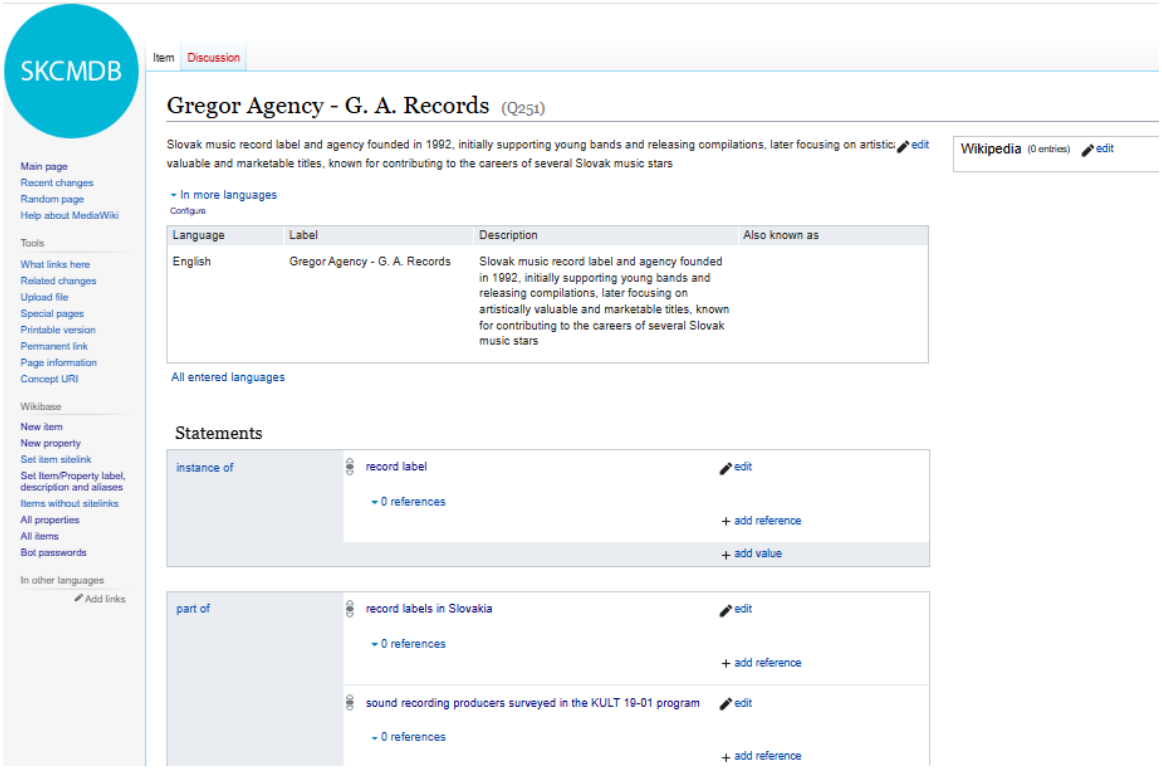
Through the Open Music Registers, we have registered music event organisers relevant to the KULT 16-01 framework, enabling researchers to distinguish which public events and venues are already statistically monitored and which are not. In cooperation with the Ministry of Culture, we have access to several years of historical microdata from the KULT 16-01 survey. This allows for building expanded statistical collections that, by combining the existing KULT 16-01 register with broader datasets from the Open Music Registers, can generate more complete, retrospectively harmonised indicators for audience development, geographic coverage of music events, and the economic structure of the live music sector. It also enables the construction of corrected, backcasted time series based on an enriched and more representative survey frame.

In addition, based on our harmonised metadata model, we can offer targeted suggestions for improving future editions of the KULT 16-01 survey, such as better distinguishing between small and large event organisers, tracking informal or semi-professional events, and improving the granularity of financial and workforce data collection.

2.5 KULT 19-01

The KULT 19-01 survey is an official statistical instrument mandated by the Ministry of Culture of the Slovak Republic and registered by the Slovak Statistical Office (Ministerstvo kultúry Slovenskej republiky 2022b). It is conducted annually under the legal framework of the State Statistics Act No. 540/2001 Coll. The survey collects detailed data on the production and distribution of sound recordings of musical works. It focuses on the volume and structure of published recordings, the

economic performance of producers and distributors, and the associated human resources. Typical questions include the number of published sound recording titles, the musical genres and types of authors involved (Slovak or foreign, soloists or ensembles), the form of distribution (physical media versus online), sales figures, revenue sources (including grants and sponsorships), and detailed costs related to production and distribution.



The G. A. Records is one of the few entries in the KULT 16-01 register that we found to be certain to have an ongoing presence in releasing sound recordings.

Through the Open Music Registers, we have registered numerous music publishers, recording companies, and independent producers active in the Slovak music market. This enables researchers to distinguish which entities are formally covered by KULT 19-01 and which operate outside the current statistical monitoring. Thanks to our cooperation with the Ministry of Culture, we also have access to historical microdata from the KULT 19-01 survey. By combining the KULT 19-01 sample frame with the broader, semantically harmonised register of music economic actors from Open Music Registers, it becomes possible to produce improved statistical collections, consistent indicators of music production and distribution over time, and corrected, retrospectively harmonised economic measurements for the recorded music sector.

However, the KULT 19-01 framework has several structural problems:

First, it blurs two distinct activities: the work of sound recording studios and record labels. Today, recording costs are dramatically lower than in the 20th century, and many bands own and operate their own recording equipment. With open European borders, sound recording often happens in a different country from where the artists reside.

More critically, Slovakia lacks a dedicated sound recording registrar. The management of ISO-standard sound recording identifiers (ISRC codes) remains anchored in Czechia, within the same organisation

that historically served Czechoslovakia. Unlike Hungary, where Mahasz acts as both a collective rights management organisation and national ISRC agency, Slovakia operates a fragmented system:

- One body registers sound recordings for commercial release (ČNS IFPI, in Czehcia.)
- Another body, SLOVGRAM, handles rights collections — requiring separate registration.

This results in significant gaps.The majority of Slovak repertoire is not registered under the SK domain of ISRC at all. Consequently, the **identification of sound recordings** and the correct **attribution of neighbouring rights** to Slovak composers and performers is an **ongoing problem**.

Our analysis—comparing an extensive collection of Slovak sound recordings, their ISRC codes, registrant domains, and the KULT 19-01 survey data—shows **minimal overlap** between these datasets.

To address this, **Open Music Registers** proposes a strategic solution:

- **Identify all sound recordings** associated with every known Slovak musical work.
- **Reconstruct a missing register** of labels and distributors connected to the Slovak repertoire.

This work is introduced here, but it will be detailed in full in **WP2**.

2.6 Composers

The Slovak Music Economy Register includes composers based on two complementary criteria: their *cultural-historical significance* and their *economic activity*. Our aim is to capture not only the heritage of Slovak musical creation but also the contemporary economic realities of active composers.

First, we include in our SKCMDb database *every composer who was born within the modern territory of the Slovak Republic* and worked here for a significant period. This anchors the register in a culturally grounded definition of Slovak music authorship. This is an important dataset for our musical work and sound recording register, but it is not directly useful for the measuring the economic activities of Slovak authors; many of these authors had long deceased and even their heirs do not receive royalties. Or some of them have left the country and they receive royalties abroad.

For the purposes of economic measurement, we include *every composer who received royalties in the previous calendar year*. This dynamic group—what we call the “active composers”—reflects the living, economically significant body of creators currently generating musical works and revenues.

One of the most authoritative sources for identifying active composers is the membership register of the Slovak collective management organisation **SOZA**. This register includes all living composers who have registered at least one musical work in Slovakia and received rights payments for at least one work during the preceding year.

Due to GDPR constraints, we cannot directly incorporate personal-level data from SOZA into the Slovak Music Economy Register. However, we carefully describe the metadata model used, and we rely on **aggregate, anonymised distributions** to guide research designs and sampling strategies.

The SOZA register is particularly valuable for constructing a **survey frame** when studying either the creation of new musical works or the revenues they generate. It captures both:

- **Composers represented by music publishers**, whose royalties often flow through business entities, and

- **Self-publishing composers**, whose incomes are collected directly and are not mediated by third-party enterprises.

In previous studies—including the *Feasibility Study on Promoting Slovak Music in Slovakia & Abroad* (Antal 2020) and the D1.1 deliverable *European-wide Methodological Assessment* (Antal, Kmety Barteková, and Remeňová 2023)—we highlighted the severe difficulty of surveying the *informal economy* in the cultural sector. Many creators operate outside formal business structures, meaning that conventional sampling frames, like business registers, systematically exclude them. This makes the SOZA dataset a uniquely crucial foundation for research, policy development, and rights administration.

To ensure that our survey samples accurately reflect income patterns—not just demographic or occupational characteristics—we employ an *iterative sampling strategy*. New respondents are sampled until the *empirical distribution of self-reported royalties* matches the *anonymised true royalty payout curve* derived from SOZA data. This method, developed in our earlier **CEEMID** and **Open Music Europe** projects, significantly improves the reliability of income distribution analyses within creative labour markets.

By including composers in the Slovak Music Economy Register—even if only at the metadata level—we build a critical bridge between *cultural identity*, *economic activity*, and *legal representation*. This integrated approach strengthens statistical validity and enhances the policy relevance of research on the Slovak music economy.

2.7 Performers

The broader SKCMDb extends beyond economic measurement to include cultural and artistic dimensions, supporting both statistical use cases (such as constructing a music satellite account for GDP calculations) and broader research and policy objectives (such as cultural diversity monitoring and repertoire promotion). SKCMDb connects datasets from rights management organisations, public archives, statistical offices, and libraries, applying a harmonised metadata model based on standards such as CIDOC-CRM, Dublin Core, and ESCO/NACE classifications.

By combining flexible, decentralised data integration with high standards of semantic interoperability, the Slovak Music Economy Register and SKCMDb represent a unique example of a sector-specific “micro data space” for the creative industries — one that is both ready for immediate use and designed for future expansion across Europe.

3 Surveying

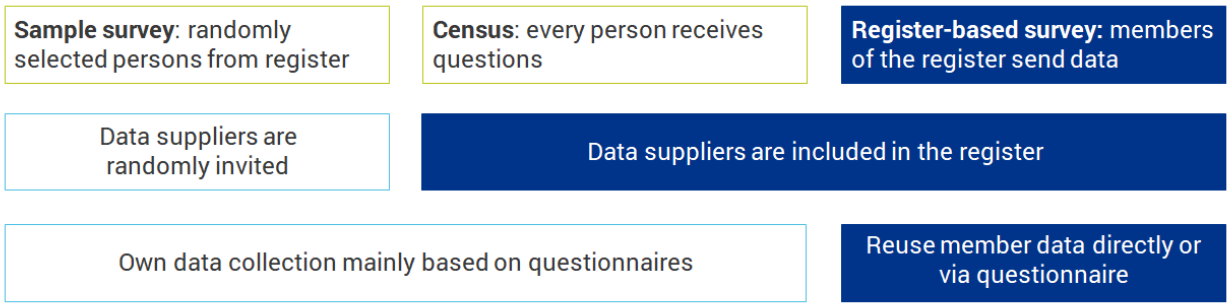
Using our register to support surveys requires integration with two important frameworks: the *Generic Statistical Information Model* (GSIM) and the *Data Documentation Initiative* (DDI). This section focuses on identifying the *universe* (the full target population) and the *coverage* (the portion of the population actually surveyed). These elements must be clearly linked to survey responses, which are then transformed into measurable variables.

Our goal is to build a harmonised approach that supports both ex-ante and ex-post survey design. A more complete technical discussion will be presented in a separate report. Here, we concentrate on ensuring that survey designs start from a well-defined universe and achieve measurable coverage.

The **Statistical Business Register (SBR)** is a foundational tool for statistical production (Eurostat 2021a). It provides the basis for constructing survey frames—deciding who is included in a data collection effort. SBRs are used to identify target populations for both sample-based and census-style surveys. Our music economy register operates in parallel to these structures, adapted to cultural sector specifics, and remains interoperable with national and European statistical systems.

Depending on the survey method, the register can support three main approaches:

- **Sample surveys**: The register serves as the sampling frame from which a representative subset of the population is randomly selected for questionnaire or interview requests.
- **Census-style surveys**: Every unit listed in the register—such as all active music publishers or labels—is contacted to supply data. The goal is full population coverage.
- **Register-based surveys**: Data is not collected through questionnaires but retrieved directly from systems of data holders (e.g., collective management organisations or tax authorities). The register ensures the completeness and traceability of linked data sources.



Comparison of three types of statistical surveys. Based on A. Wallgren and B. Wallgren: Register-based Statistics.

3.1 Extending with GSIM

To make the register usable for survey-based statistical production, it must be extended in line with the Generic Statistical Information Model (GSIM). In particular, we need to connect the identification of agents and activities to each phase of the survey lifecycle, from question design to data analysis.

In an *ex-ante* context, this means linking survey questions to structured representations: concepts, variables, classifications, and permissible value domains. This ensures that each question is grounded in the register's metadata and semantic model.

In an *ex-post* context, responses must be linked back to these same representations, so that results can be analysed in relation to registered attributes, such as legal status, sector classification, or professional role.

This alignment allows us to treat surveys and registers as components of a shared, interoperable statistical infrastructure. It enhances transparency and reusability by ensuring that both inputs (questions) and outputs (variables) are traceable to their conceptual basis and population frame.

GSIM, SDMX, and DDI are complementary standards that support this interoperability across statistical and research systems:

- *GSIM* provides the conceptual framework for describing information objects used in the production of official statistics, including surveys, registers, and outputs (United Nations Economic Commission for Europe 2025).
- *SDMX* defines data structures and exchange formats for sharing statistical data and metadata between organisations (SDMX 2021).
- *DDI* offers metadata structures for documenting survey processes—covering concepts, questions, response formats, coding schemes, and derived variables (Data Documentation Initiative 2020a).

Together, these standards enable us to harmonise metadata across systems and ensure semantic consistency from data collection through to analysis.

3.1.1 Identifying the subjects

The *Agent* classes in our model correspond to the data subjects within the target population. In administrative or enterprise surveys, we can directly link responses to agents via unique identifiers—such as a company registration number.

In population-based surveys, responses are often anonymous. In these cases, we rely on well-structured categorical variables to interpret and generalise results. These categories help profile respondents (e.g., gender, role, occupation) and ensure that response quotas are met—for example, to avoid surveys completed only by men or only by composers.

Assigning respondents to groups such as *Composers* or *Performers* relies on these pre-defined categorical variables. This structure corresponds directly to the roles and occupations recorded in the register.

Such groupings allow us to calculate indicators like the average income or number of works composed separately for performers, composers, and producers.

3.1.1.1 Occupational and economic outcomes

Often, we aim to measure the outcomes of creative and economic activities as numerical variables—for example, the number of works composed during the reference period, or the royalties received.

Temporal logic is important: royalties reported in, say, 2025 typically relate to works composed or registered earlier.

To ensure reliable measurement, numerical variables must be paired with well-defined units of measure (such as euros or forints) and logically coded missing value types:

- *Filtered*: the respondent was not asked the question.
- *Declined*: the respondent refused to answer.
- *Unknown*: the respondent did not know the answer.

These codes must be designed *ex-ante* and consistently applied in the data model.

3.1.1.2 Ex-post processing

When variables and categories are well-defined *ex-ante*, *ex-post* processing becomes straightforward.

Numerical variables can be aggregated using standard statistical measures (e.g., mean, median, weighted mean), while missing value categories are preserved and counted separately.

Groupings such as respondent country can support stratified analysis. For example, we can calculate separate income estimates for Slovakia and Hungary, and place them into an SDMX-compliant data cube, where each dimension (role, geography, year) can be subsetted or sliced for detailed analysis.

3.2 Surveying the music economy

As described in Appendix A, our register is built around *agents*—human or corporate—whose activities are recorded for economic, cultural, or statistical purposes. For survey purposes, we extend this model with two key concepts from the Data Documentation Initiative (DDI): the *universe* (the target population) and *coverage* (the actual surveyed population).

3.2.1 Open Music Europe WP1: Music Economy

In Task 1.4 and related activities, we aimed to collect more complete data about the Slovak, Hungarian, and Polish music economies. In this context, we are interested in surveying economic activity—what music business actors do, earn, and contribute.

We define *music business actors* as individuals, groups, or legal entities engaged in economically meaningful activities: earning royalties, employing others, or generating gross value added. This gross value added contributes to regional or national GDP.

These actors organise performances, sell tickets, release recordings, and license music for various uses such as recordings, film, or theatre. Their activities are classified according to:

- NACE/ISIC codes for corporate bodies (in the United States and United Kingdom: SIC), and
- ISCO/ESCO codes for individual occupations (in the United Kingdom: SOC).

Sole proprietors are a key category in music. They often have national-level five-digit NACE codes (beyond the harmonised four-digit level), reflecting their specific business activities, although these are not harmonised EU-wide. They may also be assigned ISCO/ESCO occupation codes—even if they are unaware of these classifications themselves.

3.2.2 Performing arts agents

On the performance side, the primary agents are:

- Individuals in roles such as musician, instrumentalist, or vocalist
- Corporate bodies such as music groups (bands, orchestras, ensembles, duos, trios, choirs)

These groups take diverse organisational forms. Some orchestras are formally incorporated as partnerships or limited liability companies. Others are embedded within public institutions, such as broadcasters or theatres. Many popular music groups function informally, without a registered legal entity.

While linking performing groups to business entities is possible, our register focuses on capturing the *corporate body* identity of the music group, irrespective of its legal form. On public platforms such as streaming services, the group name is usually the observable identity, not the underlying business registration.

3.2.3 Music publishing and sound recording agents

On the recording side, agents include those who:

- License works for sound recording
- Produce and distribute recordings commercially
- Facilitate public performance or synchronisation with audiovisual content

These agents include music publishers, record labels, and distributors.

Music publishers are relatively easy to register, as they are usually incorporated legal persons. Under the *Collective Rights Management Directive* (European Parliament and Europenan Council 2014), their identities must be made public if they rely on collective management. This allows direct inclusion in the register.

Record labels, by contrast, are more loosely defined. While many are legal entities, some operate informally or as subsidiaries within larger businesses. Only large, formalised labels typically have music publishing or releasing as their primary NACE-classified activity.

The national IFPI register provides a partially public resource: it reserves ISRC registrant identifiers for commercial releases. These identifiers operate similarly to ISBNs for books, enabling tracking of commercial recordings. However, ISRC registrants are not always labels:

- Sub-publishing and sub-labelling are common
- Small labels may release recordings via external distributors
- Global distributors (such as TuneCore) often act as direct ISRC registrants (ISO 2019b)

This complexity means that while ISRC codes help trace sound recordings, they do not always resolve cleanly to a label or economic agent. Our register accommodates this by recording the observable entity—such as the name under which a recording is released—and linking it, where possible, to a business or legal identity.

3.3 Surveying music diversity

3.3.1 Open Music Europe WP2: Music Diversity

Tasks T2.2 and T2.3 of the *Open Music Europe* project focus on developing better tools to measure music diversity in Slovakia, Hungary, and Poland.

Diversity in music can refer to several dimensions: the identity of creators and performers, the musical works themselves, or how works are programmed and distributed. Initiatives such as KeyChange seek measurable gender balance on stages, while funding bodies increasingly promote the visibility of women composers and ethnic minority cultures.

Another dimension of diversity is access to a broad repertoire, monitored through tools such as broadcast quotas or public performance regulations. These often include requirements for minimum percentages of local music, youth-created content, or works by women and underrepresented groups. Some diversity objectives are also genre-specific.

Capturing this range of diversity requires a register that connects people to works and works to their manifestations. For example, Slovak local content regulations define “Slovak music” by linking Slovak-language lyrics or performer/composer residence to a work or recording. Other rules, such as French or Hungarian radio quotas, similarly link agent properties—such as age, gender, or ethnicity—to musical outputs.

These requirements overlap with economic measurement but also introduce greater complexity. For corporate bodies such as music groups, diversity analysis may require details on group composition (e.g., gender, age, ethnicity)—categories considered sensitive personal data under GDPR. Capturing such data is essential for monitoring benchmarks like KeyChange but must be managed with strong safeguards.

3.3.2 Registration of musical entities

To support this kind of measurement, our data model introduces a distinct class of *musical entities*, including:

- *Musical works*
- *Sound recordings* (fixations of performances)
- *Live performances* (public-facing, non-rehearsal events)
- *Print music* (sheet music, scores)
- *Lyrics* (often literary works linked to compositions)

These entities are currently spread across fragmented, often non-public registers maintained by various rights organisations. Connecting them requires careful negotiation between transparency and proprietary control. In many cases, registration is tied to commercial incentives—such as rights management and royalty tracking—which creates tensions when such data is reused for policy evaluation or diversity monitoring.

For example, linking performer demographics to a sound recording may reveal important insights about representation on streaming platforms, but it also raises questions about data access, consent,

and equitable visibility. Navigating this space requires both legal interoperability and a trust framework between data holders.

4 Conclusion

This deliverable demonstrates that improving the economic measurement of music is not primarily a question of collecting more data, but of coordinating existing data through interoperable registers. The current fragmentation of registration processes results in duplicated effort, inconsistent identifiers, and limited statistical usability, even where relevant data is already available.

By introducing the concept of Open Music Registers, this report establishes a feasible and scalable alternative to centralised data integration. The approach builds on existing institutional data assets and enables their coordinated use through shared identifiers, harmonised classifications, and incremental alignment of data structures. Rather than requiring full standardisation or system replacement, it supports interoperability across heterogeneous data sources while preserving institutional autonomy.

The analysis identifies three core conditions for improving measurement capacity: consistent identification of agents and activities, reliable linkage between data sources, and sufficient semantic alignment across domains. These conditions are currently only partially met, which explains the limited comparability and usability of music sector statistics across Europe.

At the same time, the implementation of interoperable registers depends on addressing several structural constraints. These include aligning incentives among stakeholders with differing economic and institutional interests, ensuring compliance with legal and data protection requirements, and maintaining sufficient semantic consistency across domains such as rights management, cultural heritage, and official statistics. These are not conceptual obstacles, but coordination challenges that must be managed within existing governance frameworks.

D1.4 establishes the statistical and data-structural foundation required to address these issues. It clarifies how registers can function as the core infrastructure for economic measurement and how interoperability can be achieved without centralisation. The operational and system-level implications of this approach are further specified in D5.7, which complements this deliverable by detailing implementation pathways and integration mechanisms.

Overall, Open Music Registers provide a pragmatic framework for improving data quality, reducing administrative duplication, and enabling more reliable and comparable economic analysis of the European music sector.

© 2023 OpenMusE | HORIZON-CL2-2022-HERITAGE-01-05 | Grant Agreement No. 101095295

References

Alexiev, Vladimir, Plamen Tarkalanov, Nikola Georgiev, and Lilia Pavlova. 2020. “Bulgarian Icons in Wikidata and EDM.” *Digital Presentation and Preservation of Cultural and Scientific Heritage* 10: 45–63. https://doi.org/10.55630/dipp.2020.10.2.

Anders, Wallgren, and Wallgren Britt. 2007. *Register-Based Statistics. Administrative Data for Statistical Purposes*. 1st ed. Chichester:United Kingdom: John Wiley & Sons Ltd.

Antal, Daniel. 2020. “Feasibility Study on Promoting Slovak Music in Slovakia & Abroad.” https://doi.org/10.5281/zenodo.6427514.

———. 2023. “Pilot Program for Novel Music Industry Statistical Indicators in the Slovak Republic.” https://doi.org/10.5281/zenodo.10372026.

Antal, Daniel, Mária Kmety Barteková, and Katarína Remeňová. 2023. “Economy of music in Europe: Novel data collection methods and indicators.” Zenodo. https://doi.org/10.5281/zenodo.8334648.

Archives Expert Group on Archival Description, International Council on. 2023. “Records in Contexts–Conceptual Model. Version 1.0.” International Council on Archives. https://www.ica.org/app/uploads/2023/12/RiC-CM-1.0.pdf.

Bekiari, Chryssoula, George Bruseke, Erin Canning, Martin Doerr, Philippe Michon, Christian-Emil Ore, Stephen Stead, and Velios Athanasios, eds. 2024. “Definition of the CIDOC Conceptual Reference Model.” CIDOC CRM Special Interest Group. https://www.cidoc-crm.org/sites/default/files/cidoc_crm_version_7.2.4.pdf.

Bianchini, Carlo, Stefano Bargioni, and Camillo Carlo Pellizzari di San Girolamo. 2021. “Beyond VIAF Wikidata as a Complementary Tool for Authority Control in Libraries.” *Information Technology and Libraries* 40 (2). https://doi.org/10.6017/ital.v40i2.12959.

Bianchini, Carlo, and Lucia Sardo. 2022. “Wikidata: A New Perspective Towards Universal Bibliographic Control.” *JLIS.it* 13 (1): 291–311. https://doi.org/10.4403/jlis.it-12725.

“Business-to-Government Data Sharing for Public Interests in the European Union: Results of a Public Consultation.” 2022. Lecture Notes in Computer Science. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-15086-9_33.

Camp, Ann Van, Sven Lieber, and IFLA. 2022. “ISNI, a Top Tool for Quality Enhancement, Smooth Data Flows and Efficient Internal Processes.” Dublin: Ireland: International Federation of Library Associations; Institutions (IFLA). https://repository.ifla.org/handle/123456789/2008.

Chartered Accountants in England, The Institute of, and Wales. 2023. “Appendix 2: Example Statements for Inclusion in the Accounts.” https://www.icaew.com/technical/corporate-reporting/uk-gaap/uk-gaap-faqs/micro-entities-filing-options#subheading-c2ccbfd5-c584-4185-931c-68cc40c8eada.

CIDOC CRM Special Interest Group. 2025. *CIDOC Conceptual Reference Model (ISO 21127:2024)* (version 7.1.3). International Council of Museums (ICOM). https://cidoc-crm.org/versions-of-the-cidoc-crm.

Commission, European. 2017. “ESCO Handbook.” https://doi.org/10.2767/934956.

Core, Dublin. 2020. “DCMI Metadata Terms.” http://dublincore.org/specifications/dublin-core/dcmi-terms/2020-01-20/.

Curry, Edward. 2020. “Dataspaces: Fundamentals, Principles, and Techniques.” In *Real-Time Linked Dataspaces: Enabling Data Ecosystems for Intelligent Systems*, 45–62. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-29665-0_3.

Data Documentation Initiative. 2020a. *DDI Codebook 2.5 Documentation* (version 2.5). DDI Alliance. https://ddialliance.org/Specification/DDI-Codebook/2.5/.

———. 2020b. “DDI Lifecycle (3.3) Documentation.” https://ddi-lifecycle-documentation.readthedocs.io/en/latest/index.html.

———. 2020c. *DDI Representation Types Vocabulary*. DDI Alliance. https://ddialliance.org/vocabulary/representationtype.

Data Documentation Initiative Alliance. 2025. *DDI-RDF Discovery Vocabulary (Disco)* (version 1.1). DDI Alliance. https://rdf-vocabulary.ddialliance.org/discovery.html.

De Smedt, Johan. 2023. “The ESCO Ontology. Version 1.2.” European Commission. http://data.europa.eu/esco/model/1.2.

Diefenbach, Dennis, Max de Wilde, and Samantha Alipio. 2021. “Wikibase as an Infrastructure for Knowledge Graphs: the EU Knowledge Graph.” In *ISWC 2021*, 12922:631–47. Online, France: Springer. https://doi.org/https://doi.org/10.1007/978-3-030-88361-4_37.

Directorate-General for Research and Innovation (European Commission). 2022. “Report on a European Collaborative Cloud for Cultural Heritage: Ex-Ante Impact Assessment.” European Commission. https://data.europa.eu/doi/10.2777/64014.

Dublin Core Metadata Initiative. 2025. *DCMI Metadata Terms* (version 2020-01-20). Dublin Core Metadata Initiative. https://www.dublincore.org/specifications/dublin-core/dcmi-terms/.

EBU, and Gaia-X. 2022. “Dataspace for Cultural and Creative Industries. Position Paper. v.2.0.” Gaia-X. https://gaia-x.eu/wp-content/uploads/2022/10/EBU_position-paper_Media-Data-Space.pdf.

Elias, Peter, Andy Dickerson, and Neil Bachelor. 2023. “A Skills Classification for the UK: Plans for Development and Maintenance.” Government Social Research. https://assets.publishing.service.gov.uk/media/652fdb9d92895c0010dcb9a5/A_skills_classification_for_the_UK.pdf.

ESS. 2017. “Position Paper on Access to Privately Held Data Which Are of Public Interest.” European Statistical System. https://ec.europa.eu/eurostat/documents/13019146/13346094/ESS+Position+Paper+on+Access+to+privately+held+data+final+-+Nov+2017.pdf/6ef6398f-6580-4731-86ab-9d9d015d15ae?t=1511447619000.

ESTAT. 2022. “Statistical Classification of Economic Activities in the European Community, Rev. 2.1.” https://showvoc.op.europa.eu/#/datasets/ESTAT_Statistical_Classification_of_Economic_Activities_in_the_European_Community_Rev._2.1._%28NACE_2.1%29/data.

European Commission. 2020a. “ESCO Ontology.” European Commission. https://data.europa.eu/esco/model.

———. 2020b. “A European Strategy for Data. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Region.” European Commission. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0066.

———. 2021. “Commission Recommendation (EU) 2021/1970 of 10 November 2021 on a Common European Data Space for Cultural Heritage.” European Commission. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32021H1970.

European Commission, and Directorate-General for Digital Services. 2017. *New European Interoperability Framework – Promoting Seamless Services and Data Flows for European Public Administrations*. Publications Office. https://doi.org/doi/10.2799/78681.

European Parliament and Council of the European Union. 2022. “Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European Data Governance and Amending Regulation (EU) 2018/1724 (Data Governance Act).” *Official Journal of the European Union* OJ L (152): 1–44. https://data.europa.eu/eli/reg/2022/868/oj.

———. 2023. “Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on Harmonised Rules on Fair Access to and Use of Data and Amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act).” *Official Journal of the European Union* OJ L (2023/2854): 1–48. https://data.europa.eu/eli/reg/2023/2854/oj.

European Parliament, and Europenan Council. 2014. *Directive 2014/26/EU of the European Parliament and of the Council of 26 February 2014 on Collective Management of Copyright and Related Rights and Multi-Territorial Licensing of Rights in Musical Works for Online Use in the Internal Market*. European Union. https://eur-lex.europa.eu/eli/dir/2014/26/oj.

Europeana. 2017. “Definition of the Europeana Data Model V5.2.8.” Europeana. https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation//EDM_Definition_v5.2.8_102017.pdf.

Eurostat. 2008. “NACE Rev. Statistical Classification of Economic Activitiesin the European Community.” Methodologies and Working Papers. Office for Official Publications of the European Communities. https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF.pdf/dd5443f5-b886-40e4-920d-9df03590ff91?t=1414781457000.

Eurostat. 2021a. “European Business Statistics Methodological Manual for Statistical Business Registers.” Eurostat. https://ec.europa.eu/eurostat/documents/3859598/11029628/KS-GQ-20-010-EN-N.pdf.

———. 2021b. *European Business Statistics Methodological Manual for Statistical Business Registers. 2021 Edition*. Manuals and Guidelines. Luxembourg: Publications Office of the European Union. https://doi.org/10.2785/093371.

———. 2024. “NACE Rev. 2.1 Linked Data Representation.” European Commission. https://showvoc.op.europa.eu/#/datasets/nace/data.

Fagerving, Alicia. 2023. “Wikidata for Authority Control: Sharing Museum Knowledge with the World.” *Digital Humanities in the Nordic and Baltic Countries Publications* 5 (1): 222–39. https://doi.org/10.5617/dhnbpub.10665.

Fischer, Barbara, Joel Thill, and Anila Angjeli. 2020. “Providing Sustainable Data Services Through Wikibase and Wikidata.” Presentation at SEMIC 2020 Conference. https://joinup.ec.europa.eu/sites/default/files/custom-page/attachment/2020-11/Parallel-track-4_B-Fischer_J-Thill_A-Angjeli%20final%20ppt.pdf.

Gilles, Mathieu. 2023. “A General Presentation of the European Open Science Cloud (EOSC).” Zenodo. https://doi.org/10.5281/zenodo.10228651.

Green, Lucy. 2017. *Music, Informal Learning and the School: A New Classroom Pedagogy*. London: Routledge. https://doi.org/10.4324/9781315248523.

Group, DataCite Metadata Working. 2021. “DataCite Metadata Schema 4.4.” DataCite e.V. https://schema.datacite.org/meta/kernel-4.4/.

International Federation of the Phonographic Industry (IFPI). 2021. *ISRC Handbook: The International Standard Recording Code*. 4th Edition. IFPI. https://www.ifpi.org/wp-content/uploads/2021/02/ISRC_Handbook.pdf.

International Labour Office. 2015. “Recommendation R204 - Transition from the Informal to the Formal Economy Recommendation, 2015 (No. 204).” International Labour Office. https://www.ilo.org/dyn/normlex/en/f?p=NORMLEXPUB:12100:0::NO::P12100_ILO_CODE:R204.

International Labour Organization. 2024. “ISCO Linked Data Representation.” International Labour Organization. https://ilo.org/isco.

ISO. 2017a. “ISO/IEC 19941:2017(en), Information Technology — Cloud Computing — Interoperability and Portability.” London:United Kingdom: International Organization for Standardization. https://www.iso.org/obp/ui/#iso:std:iso-iec:19941:ed-1:v1:en.

———. 2017b. “ISO/IEC 5127:2017(en), Information and Documentation — Foundation and Vocabulary.” London:United Kingdom: International Organization for Standardization. https://www.iso.org/obp/ui/en/#iso:std:iso:5127:ed-2:v1:en.

———. 2019a. “ISO/IEC 20546:2019 Information Technology — Big Data — Overview and Vocabulary.” London:United Kingdom: International Standards Organisation. https://www.iso.org/obp/ui/en/#iso:std:iso-iec:20546:ed-1:v1:en.

———. 2019b. “International Standard Recording Code (ISRC). ISO 3901:2019.” International Organization for Standardization. https://www.iso.org/standard/64817.html.

———. 2021. “ISO 10957:2021 - Information and Documentation — International Standard Music Number (ISMN).” International Organization for Standardization. https://www.iso.org/standard/83122.html.

———. 2022. “International Standard Musical Work Code (ISWC). ISO 15707:2022.” International Organization for Standardization. https://www.iso.org/standard/83125.html.

———. 2023. “ISO/IEC 2382:2015(en), Information Technology — Vocabulary.” London:United Kingdom: International Standards Organisation. https://www.iso.org/obp/ui/en/#iso:std:iso-iec:2382:ed-1:v2:en.

Lebo, "Timothy, Satya Sahoo, Deborah McGuinness, Khalid Belhajjame, James Cheney, David Corsar, Daniel Garijo, Stian Soiland-Reyes, Stephan Zednik, and Jun Zhao". "2013". “"PROV-o: The PROV

Ontology".” "W3C Recommendation". "World Wide Web Consortium". https://www.w3.org/TR/2013/REC-prov-o-20130430/.

Manzoni, Laura. 2022. “Identifiers Within Conceptual Models and Standards.” *JLIS.it* 13 (3): 91–103. https://doi.org/10.36253/jlis.it-473.

Ministerstvo kultúry Slovenskej republiky. 2022a. “KULT (MK SR) 16—01 Ročný Výkaz o Verejných Podujatiach v Oblasti Profesionálnej Hudobnej Kultúry.” Ministerstvo kultúry Slovenskej republiky. https://www.culture.gov.sk/wp-content/uploads/2021/08/KULT-16-01_o_verejnych_podujatiach_v_oblasti_profesion_hudob_kultury_2022.pdf.

———. 2022b. “KULT (MK SR) 19—01 Ročný Výkaz o Výrobe a Distribúcii Zvukových Záznamov Hudobných Diel.” Ministerstvo kultúry Slovenskej republiky. https://www.culture.gov.sk/wp-content/uploads/2021/08/KULT-19-01_o_vyrobe_a_distribucii_zvukovych_zaznamov_hudobnych_diel_2022.pdf.

———. 2022c. “KULT (MK SR) 5—01 Ročný Výkaz o Hudobnom Telese a Umeleckom Súbore.” Ministerstvo kultúry Slovenskej republiky. https://www.culture.gov.sk/wp-content/uploads/2021/08/KULT-5-01_o_hudobnom_telese_a_umeleckom_subore_2022.pdf.

Nagel, Lars, and Douwe Lycklama, eds. 2021. “Design Principles for Data Spaces. Position Paper. Version 1.0.” Open DEI. https://doi.org/10.5281/zenodo.5244997.

Národná rada Slovenskej republiky. 2001. “Zákon č. 540/2001 z. Z. O Štátnej Štatistike.” Národná rada Slovenskej republiky. https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2001/540/20220101.

National Statistics, The Office of. n.d. “Standard Industrial Classification.” https://classification.codes/classifications/industry/uk-sic.

Open Music Europe. 2023. “Open Music Europe (OpenMusE) – An Open, Scalable, Data-to-Policy Pipeline for European Music Ecosystems.” https://doi.org/10.3030/101095295.

Pfeiffer, Michelle, and Jose Emilio Labra Gayo. 2021. “Representing CIDOC-CRM in Wikibase for the Luxembourg Shared Authority File.” https://swib.org/swib21/slides/05-03-gayo.pdf.

Rossenova, Lozana, Paul Duchesne, and Ina Blümel. 2022. “Wikidata and Wikibase as Complementary Research Data Management Services for Cultural Heritage Data.” In *CEUR Workshop Proceedings*. https://serwiss.bib.hs-hannover.de/frontdoor/deliver/index/docId/2573/file/rossenova_etal2022-wikidata_research_data_mgmt.pdf.

Sæbø, Hans Viggo, and Xeni Kristine Dimakos. 2023. “Use of New Data Sources – the Case of Statistics Norway.” In, 12. Ottawa:Canada: International Statistical Institute. https://www.cambridge.org/core/services/aop-cambridge-core/content/view/C739E1DE223088FD3D761466DCDA2EFE/S1867299X1800034Xa.pdf/big-data-for-good-unlocking-privately-held-data-to-the-benefit-of-the-many.pdf.

Sardo, Lucia, and Carlo Bianchini. 2022. “Wikidata: A New Perspective Towards Universal Bibliographic Control.” *JLIS.it : Italian Journal of Library and Information Science* 13 (1): 291–311. https://doi.org/10.4403/jlis.it-12725.

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.

SDMX Technical Working Group. 2021. *SDMX Information Model* (version 3.0). SDMX Sponsors. https://sdmx.org/standards-2/.

Sun, Chang. 2021. "Knowledge Graph for Microdata of Statistics Netherlands." *arXiv Preprint*. https://arxiv.org/abs/2101.07622.

"Towards a European Strategy on Business-to-Government Data Sharing for the Public Interest. Final Report Prepared by the High-Level Expert Group on Business-to-Government Data Sharing." 2020. Luxembourg: Publications Office of the European Union. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=64954.

UK Office for National Statistics. 2022. "UK Standard Industrial Classification (SIC) as Linked Data." UK Office for National Statistics. https://classification.codes/classifications/industry/uk-sic.

UNECE. 2015. "Guidelines on Statistical Business Registers." United Nations Economic Commission for Europe. http://www.unece.org:8080/fileadmin/DAM/stats/publications/2015/ECE_CES_39_WEB.pdf.

United Nations Department of Economic and Social Affairs. 2024. *Guidelines on Statistical Business Registers*. Studies in Methods, Series f 122. New York: United Nations.

United Nations Economic Commission for Europe. 2025. "Generic Statistical Information Model. UNECE Statswiki." United Nations Economic Commission for Europe. April 16, 2025. https://unece.github.io/GSIM-2.0/GSIMv2.html.

United Nations Statistics Division. 2020. "United Nations Guidelines on Statistical Business Registers. Final Draft Prior to Official Editing." United Nations. https://unstats.un.org/unsd/business-stat/SBR/Documents/UN_Guidelines_on_SBR.pdf.

———. 2024. "Guidelines on Statistical Business Registers." United Nations. https://unstats.un.org/unsd/business-stat/SBR/Documents/UN_Guidelines_on_SBR.pdf.

van Veen, Theo. 2019. "Wikidata: From 'an' Identifier to 'the' Identifier." *Information Technology and Libraries* 38 (2): 72–81. https://doi.org/10.6017/ital.v38i2.10886.

Vigorito, Andrea. 2022. "Government Access to Privately-Held Data: Business-to-Government Data Sharing: Voluntary and Mandatory Models." *European Journal of Comparative Law and Governance*. https://doi.org/10.1163/22134514-bja10030.

Wikimedia Deutschland. 2025. "Wikibase Data Model." Wikimedia Foundation. https://www.mediawiki.org/wiki/Wikibase/DataModel.

Appendix A — Open Music Economy Register

A.1 Value of the Data

Music registers are necessary for correctly administrating copyrights and their neighbouring rights, accounting for music on corporate and national levels, and surveying business or artistic activities. This database and its documentation show a novel way of creating a multi-purpose superregister that connects private and public registers.

A.2 Background

Economic measurement and rights administration in the music sector are hampered by fragmented and uncoordinated registers. These issues affect both macro-level statistics — such as national accounts and cultural satellite accounts — and micro-level tracking of royalties for authors, performers, and producers. *Open Music Europe* aims to address these challenges by introducing interoperable, standards-based registers that improve data collection and reuse.

The Slovak Music Economy Register, developed under Task T1.4 of the Open Music Europe project, builds on the statistical and methodological framing developed under that initiative. The project identified persistent weaknesses in the way Europe collects and categorises data about the music economy. Despite the proliferation of digital metadata, key economic indicators — including employment, income distribution, and investment in music production — remain poorly measured in many national contexts.

The Slovak Music Economy Register is one component of the broader Slovak Comprehensive Music Database (SKCMDb). While the economy register focuses specifically on music-related economic actors — including natural persons (e.g. composers, performers) and legal entities (e.g. publishers, record labels, venues) — the full SKCMDb also includes complementary datasets on musical works, sound recordings, and physical venues. Together, these interoperable components form a multi-purpose, standards-aligned platform for supporting research, rights management, policy, and cultural documentation in the Slovak music sector.

We strive to build a best policy practice in Slovakia within the context of the European Data Strategy (European Commission 2020b), ensuring that this register supports both national implementation and broader European data governance goals.

Two earlier deliverables of the Open Music Europe project helped define the strategy for addressing these issues:

- **A European-wide methodological assessment** (Antal, Kmety Barteková, and Remeňová 2023), delivered as D1.1 and presented to policy stakeholders in Bratislava in September 2023, highlighted how fragmented data collection limits the practical application of classifications such as NACE and ISIC. It recommended building open, artist-facing infrastructure that would allow professional metadata to be registered once, then reused across policy, research, and rights management applications.
- **Pilot Program for Novel Music Industry Statistical Indicators in the Slovak Republic** (Antal 2023) demonstrated how coordinated data flows between music organisations, registers, and national statistics authorities could be implemented in practice. It reused earlier national

surveys — KULT05, KULT16, and KULT19 — which documented creators' economic activities and institutional affiliations, but lacked persistent identifiers and update mechanisms1.

The Slovak pilot focused in part on the reuse of these earlier national surveys. These survey-based registers, while useful, lacked standardised identifiers and mechanisms for longitudinal updates. Open Music Registers aims to absorb and maintain these sources, offering a persistent, updateable register model that respects the structure of the original data while supporting modern interoperability.

The register of SOZA, the Slovak collective management organisation, will also be integrated into the platform as a core node. Together, these sources allow Open Music Registers to support both institutional and self-registered entries, bridging the gap between statistical, cultural, and operational metadata in the music economy.

A register aims to be a complete list of the objects in a specific group or population (Anders and Britt 2007). Norway created statistical registers to tap into governmental data stores in 1990 and into municipal ones in 1995; by 2019, it utilised about 100 records and drew data from 30 public institutional sources (Sæbø and Dimakos 2023, 1). Like most countries, such "administrative data" was retrieved from other governmental entities, not the private sector.

Collecting data from people, companies, and non-profits still relies on census-like comprehensive and sample surveys that take the form of filling out a digital or paper format or answering questions to an interviewer who fills out the form instead of the respondent. Surveying is costly and often inaccurate. Asking randomly selected music creators about their received royalties over a year requires the respondent to answer questions after opening and reviewing various royalty statements — or relies on fallible memory and cognitive shortcuts.

Statistics Norway realised that instead of asking 7,000 households what they were buying in the supermarket, it is far more accurate and potentially cheaper to acquire the data directly from the sales logs of the supermarkets or the payment transaction records of the credit card companies. As explained in T1.1 (Antal 2023), this model improves both accuracy and cost-effectiveness.

CEEMID, the predecessor of Open Music Europe, relied on similar techniques used by experimenting government statisticians. While continuing to anonymously survey randomly selected music creators, it also compared those results with actual anonymised payouts from Artisjus in Hungary and SOZA in Slovakia. This triangulation helped calibrate survey accuracy and validate register-based approaches.

Norway's approach is especially relevant in light of its new statistical law (in force since 2021), which permits register-based data collection following a cost/benefit and risk analysis conducted by Statistics Norway. Norway, like all EFTA countries and EU member states, participates in the European Statistical System and applies the same EU/EEA regulations that govern Slovakia, Bulgaria, and Hungary. While Open Music Europe does not propose or support compulsory state-mandated data collection, it encourages voluntary, mutually beneficial data sharing with public authorities in the music sector.

On the level of data coordination, this approach follows recent international guidelines on using privately held data for official statistics (UNECE 2015; United Nations Statistics Division 2020; Eurostat

1 The three KULT surveys (Ministerstvo kultúry Slovenskej republiky 2022c, 2022a, 2022b) are mandated on the basis of (Národná rada Slovenskej republiky 2001).

2021b), and draws more broadly from the policy and research literature on business-to-government data sharing (ESS 2017; “Towards a European Strategy on Business-to-Government Data Sharing for the Public Interest. Final Report Prepared by the High-Level Expert Group on Business-to-Government Data Sharing” 2020; Vigorito 2022; “Business-to-Government Data Sharing for Public Interests in the European Union: Results of a Public Consultation” 2022).

From a statistical point of view, our planned music industry registers can be seen as “administrative registers,” even though they were not created by a national statistical authority. They are designed for reuse in surveys and other statistical activities. See more about their application in Chapter 3.

Private institutions — such as collective management organisations, music publishers, and event promoters — often maintain far more granular and up-to-date registers than national statistical authorities. However, these internal registers are typically developed for operational purposes and are not always harmonised, standardised, or subject to the quality control procedures necessary for statistical use. They may be rich in detail but lack consistent classification, completeness, or interoperability.

To address this, we begin by harmonising public registers — which benefit from more robust methodological foundations — using international standards such as SDMX, ESCO, NACE, and ISCO (SDMX Technical Working Group 2021; European Commission 2020a; International Labour Organization 2024), and when applicable, their UK variants (UK Office for National Statistics 2022). On this basis, we construct interoperable registers of authors, music performers, and music-related businesses (covering areas like recording, publishing, and live performance). These core registers are cross-checked and enriched using data from public library systems and cultural heritage repositories, drawing on ontological patterns from DCTERMS and CIDOC CRM (Dublin Core Metadata Initiative 2025; CIDOC CRM Special Interest Group 2025).

This approach supports the harmonisation of authority files across institutional domains. It not only improves semantic alignment but also ensures traceability, transparency, and completeness in multi-use contexts — from copyright management to statistical reporting. As part of this harmonisation, we integrate both legal identifiers (e.g. national business registry numbers, OpenCorporates URIs) and creative or bibliographic identifiers such as ISNI (for creators and performers), VIAF (for authority control), ISRC (for recordings), and ISWC (for musical works). These external identifiers act as bridges, linking our internal registers to the wider ecosystem of open and authoritative data.

To complement our work with registers, we also integrate survey-based data into the same harmonised semantic framework. In our Wikibase environment, we incorporate survey questions aligned with international statistical programmes, such as the European Labour Force Survey (EU-LFS) and the Eurobarometer. By reusing harmonised question formulations and structures, we ensure both comparability and interoperability with established European data infrastructures.

Crucially, we enrich these survey elements with DDI-Discovery (DDI Codebook) metadata standards, enabling structured, machine-readable descriptions of questions, variables, and datasets (Data Documentation Initiative Alliance 2025). While the DDI metadata model is not natively available in RDF, we extend our Wikibase data model by developing a custom representation of core DDI terminology, including representation types and base types (Wikimedia Deutschland 2025; Data Documentation Initiative 2020a, 2020c). This allows us to model variables and responses with the same semantic precision we apply to entities like authors, works, or enterprises.

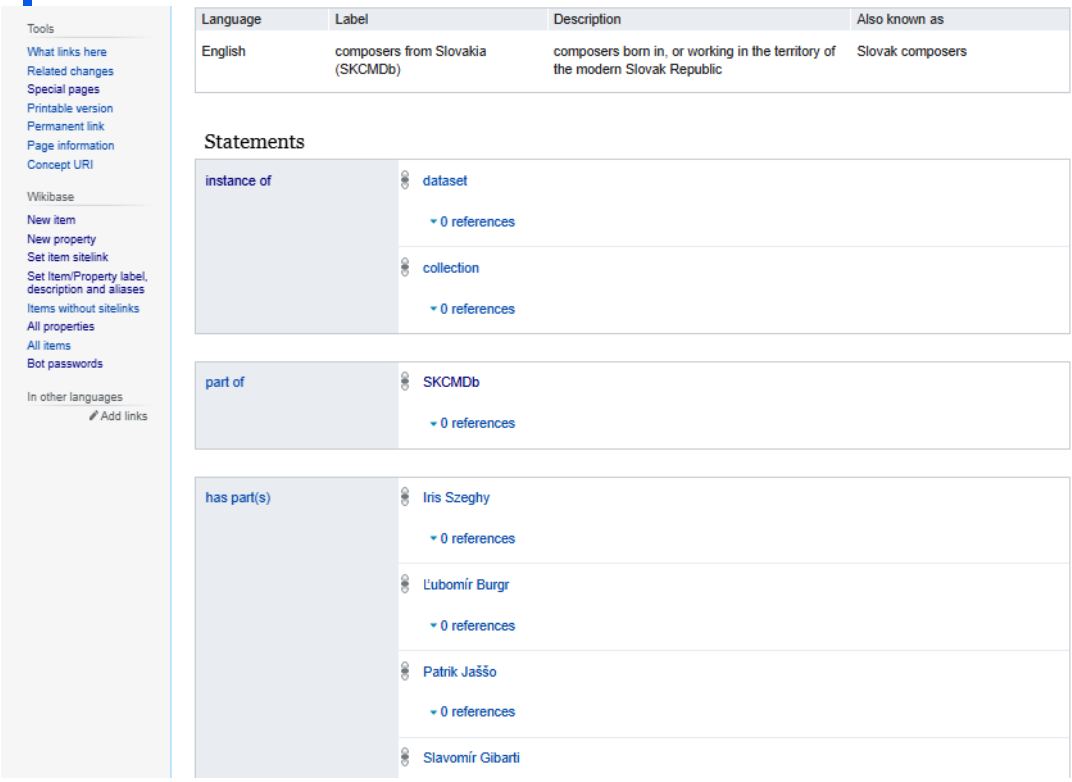
By aligning survey metadata with our broader linked data framework, we make it possible to cross-reference survey variables with register-based indicators and other statistical data — a crucial step toward integrated, multi-source, and semantically rich cultural statistics. This approach lays the foundation for automated documentation, discovery, and comparative analysis across administrative, heritage, and survey-based data sources.

A.3 Collection statement

The Slovak Music Economy Register currently includes all music publishers active in the Slovak Republic — whether registered locally or operating cross-border — as well as the full populations captured in the national cultural employment surveys KULT05, KULT16, and KULT19.

The register covers key music-sector populations: *music authors*, *music performers*, *groups of performers* (including ensembles and orchestras), *record labels* (formal or informal businesses), and *music publishers* (registered enterprises). These entities are linked to *music works* and *sound recordings*, understood as cultural products.

The register is developed to support data validation and update processes at SOZA and the Slovak Music Centre, and to serve as a statistical base for constructing a satellite business account for the Slovak music industry.



Composers:

- ☒ Every composer who was born on the territory of the modern Slovak Republic and worked here for a significant timespan.
- ☒ Every composer who received royalties in the previous calendar year (“active composers”).

The Slovak Music Economy Register includes composers based on both cultural-historical and economic-function criteria. However, one of the most authoritative sources for identifying active composers is the membership register of the Slovak collective management organisation **SOZA**. This register includes all living composers who have registered at least one musical work in Slovakia and received payment for at least one work in the preceding year.

Due to GDPR constraints, we cannot directly ingest personal-level data from SOZA into the Slovak Music Economy Register. However, we describe the metadata model in detail and use aggregate, anonymised distributions to guide our research designs.

This register is especially valuable for constructing a **survey frame** when studying the creation of new musical works or the revenue generated by them. It covers both: - **Composers represented by music publishers**, whose revenue is often routed through business entities, and - **Self-publishing composers**, whose income is not mediated by third-party enterprises but is still traceable via SOZA's collective rights distributions.

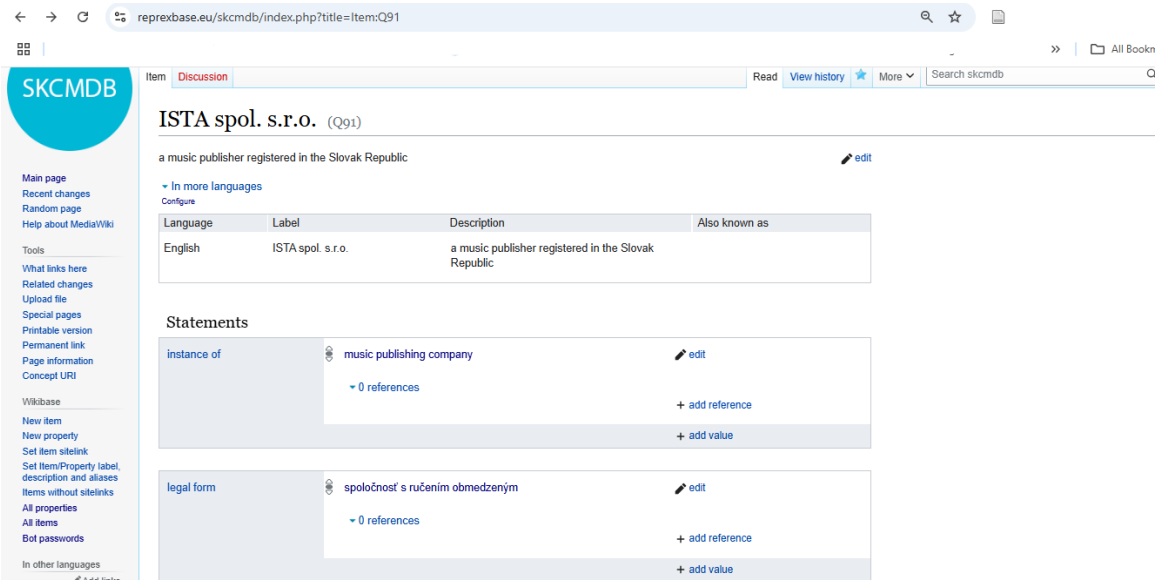
In previous work — notably the *Feasibility Study on Promoting Slovak Music in Slovakia & Abroad*, on which WP1 and WP2 are partly based (Antal 2020), and the subsequent D1.1 deliverable of WP1 *European-wide methodological assessment* (Antal, Kmety Barteková, and Remeňová 2023) — we highlighted the extreme difficulty of surveying the *informal economy* in cultural sectors. Many creators operate without formal business structures, and traditional sampling frames (like business registers) exclude them entirely. This makes the SOZA register uniquely important: it offers a grounded, operational definition of the composer population for the purposes of research, policy, and rights administration.

To ensure our survey samples are not only representative by demographic or occupational strata but also reflective of actual income patterns, we iteratively sample new respondents until the *empirical distribution of self-reported royalties* aligns with the *true anonymised royalty payout curve* from SOZA. This method, pioneered in our earlier CEEMID and Open Music Europe work, enables more reliable analysis of income distributions in creative labor markets.

The inclusion of composers in the Slovak Music Economy Register, even at the metadata level, helps bridge cultural identity, economic activity, and legal representation in a way that supports statistical integrity and policy utility.

Publishers:

- ☒ Every music publisher that was ever registered on the the territory of the modern Slovak Republic.
- ☒ Every music publisher who registered with SOZA and received royalties in the previous calendar year. ("Active music publishers".)



The entry of ISTA spol. s.r.o. includes its registration number in the official Slovak business register and in OpenCorporates

In statistical terms, we are identifying incorporated business enterprises engaged in music publishing. However, this activity is not adequately distinguished at the four-digit level of the NACE economic activity taxonomy. Instead, under a distinct legal mandate, the authors' collective management organisation (SOZA) maintains a register of active publishers — those who have received royalty payments — providing a more functional and usage-based definition.

As a result, our register may include music publishers who are not classified as such by the Slovak statistical authority. The advantage of this approach is broader visibility: we identify more entities actually engaging in music publishing. The downside is reduced specificity — our dataset may include organisations such as advertising agencies or non-profits, whose primary activity is not publishing but who receive royalties for music rights.

OpenCorporates is the world's largest open database of companies, offering structured, machine-readable data on more than 200 million companies across over 140 jurisdictions. It aggregates official records — including registration numbers, legal status, directors, and filings — directly from government sources and assigns each company a persistent identifier. These identifiers enable transparency, accountability, and seamless cross-border data linkage.

Whenever possible, we use OpenCorporates IDs — derived directly from the Slovak business register — to link our entries. This ensures that the Slovak Music Economy Register is interoperable with similar registers in other countries that are also indexed in OpenCorporates.

Producers:

- ☒ very record label that registered sound recordings under the SK top-level ISRC domain.
- ☐ very record label that registered sound recordings in the SLOVGRAM database using non-SK ISRC codes, with a known location or seat in the Slovak Republic. This group overlaps with the target population of the **KULT 19 — Production and Distribution of Sound Recordings** survey.
- ☐ Every record label that was active in the SLOVGRAM database during the previous business year (*i.e.*, registered at least one recording). These are considered *currently active music labels*.

The Slovak Music Economy Register includes and extends the frame of the KULT 19 survey. KULT 19 is an annual statistical report collected by the Ministry of Culture from producers of musical sound recordings (Ministerstvo kultúry Slovenskej republiky 2022b). It gathers data on the number and type of published titles (by genre and author origin), distribution formats (physical and online), and associated business activity — including income from sales, grants, and rights, as well as operational and labour costs. Its target population includes legal entities formally engaged in the production and distribution of music recordings in the Slovak Republic.

While our register already includes all entities covered by KULT 19, it goes further by incorporating: - Labels that operate informally or without formal NACE classification, - Cross-border entities regularly distributing music in Slovakia, - Producers whose publishing activity is captured through rights systems (like ISRC and SLOVGRAM), not just surveys.

By integrating multiple sources and maintaining persistent identifiers, the Slovak Music Economy Register supports not only statistical reporting, but also **licensing transparency, rights management, and research**. It provides a stable population frame for surveys and policy use, while remaining flexible and open to decentralised updates from the music sector itself.

Performers:

- ☒ very musical group invited into the KULT 05 statistical survey.
- ☐ Every musical group associated with Slovak labels.
- ☐ Every music performer associated with the commercially released recording of a Slovak composer's work.

At present, there is no authoritative or comprehensive register of music performers in Slovakia. The Slovak Music Economy Register incorporates performers only through indirect or partial sources.

One foundation is the KULT 05 statistical survey of musical ensembles and artistic groups, which provides a structured and repeated sample of formally recognized performing entities (Ministerstvo kultúry Slovenskej republiky 2022c). Beyond this, we attempt to expand the coverage using metadata from Slovak record labels and from recordings featuring works by Slovak composers. These entries are typically inferred via performance credits in commercial catalogues or rights databases, which may include ensemble or individual-level information.

Within the Open Music Europe project, we developed and tested **big data sampling methods** to identify performers across public platforms and streaming services, linking them to works, recordings, and labels. However, these approaches remain experimental and fragmentary. As of now, the register does not provide a complete or verified performer population.

To build a comprehensive performer register, we are working toward constructing a complete catalogue of sound recordings released by Slovak artists. This effort is underway in WP2 of Open Music Europe. Every commercially released recording is expected to carry an ISRC (International Standard Recording Code), governed by ISO 3901 (ISO 2019b). However, the ISRC system only reliably identifies the *featured artist* — the primary credited performer or ensemble — as defined in the ISRC Handbook (International Federation of the Phonographic Industry (IFPI) 2021). Session musicians and co-performers are generally excluded from this metadata.

Even when performer names are attached, legal and ethical constraints apply. If the performer is a corporate body, such as a registered ensemble or group, they may be represented in the register as

an organisation. If the performer is a **natural person**, however, the General Data Protection Regulation (GDPR) applies. This restricts our ability to publish or process personal data without explicit consent, or without a clear public-interest justification (European Commission and Directorate-General for Digital Services 2017).

In some cases — such as artist profiles hosted by the Slovak Music Centre or participation in the SKCMDb — performers have voluntarily provided such consent. As opt-in participation increases, particularly among SOZA members, our coverage will improve.

Nonetheless, we believe that connecting performers with the recordings to which they have contributed is not only possible but necessary. Performers hold a **moral right to be credited** as co-creators of the intellectual property embedded in a sound recording. The Slovak Music Economy Register is designed to evolve toward a compliant and ethically grounded model of performer inclusion.

Venues:

- ☑ Every music venue listed in the Slovak Music Centre database.

The Slovak Music Economy Register currently includes venues as cultural infrastructure nodes, primarily derived from curated databases such as the Slovak Music Centre. While there is no standardised legal definition of a music venue in Slovakia, we draw on operational data to approximate venue activity and significance.

This register aligns in part with the frame of **KULT 16 — Public Events in Professional Music Culture**, a statistical survey that collects data from organisations organising public concerts and performances. While the survey does not define venues directly, it identifies regular performance sites through data on event frequency, genre, audience capacity, and regional distribution. This provides an empirical basis for identifying key music venues indirectly.

Looking forward, we aim to enhance venue identification using authoritative geospatial infrastructure — including the *ZBGIS* national geoinformation system and the *Real Estate Cadastre* (Kataster nehnuteľností), accessible via geoportal.sk and katasterportal.sk. By linking venue records to these official GIS and address datasets, we can assign persistent, interoperable identifiers and support cross-domain use cases — from cultural policy to urban planning and sustainability research.

A.4 Data description

ⓘ

The Slovak Music Economy Register builds on well-established data models used by music agencies and rights management organisations to represent both natural persons and corporate bodies — including formal enterprises, sole proprietors, informal groups, and dissolved entities. To ensure accuracy and legal traceability, the model incorporates time and location attributes critical for capturing life events such as birth, death, inheritance, and organisational transitions like succession or dissolution. The register is partly event-based, enabling longitudinal tracking of identities, roles, and affiliations over time.

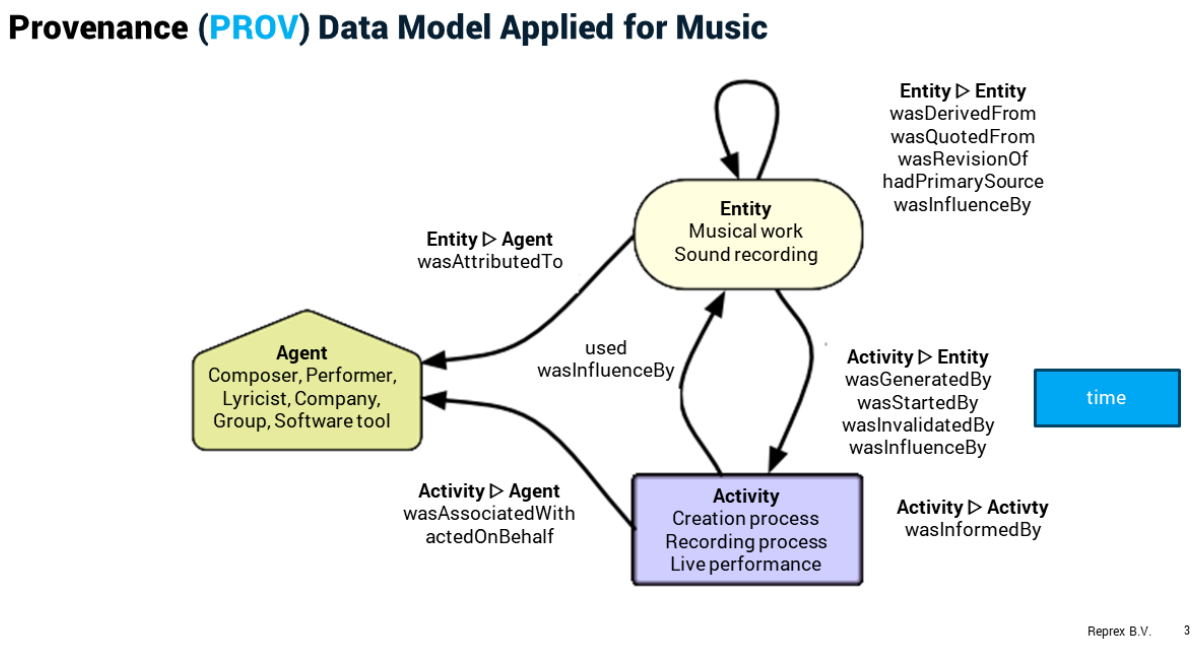
A.4.1 General concepts of surveying

Class	Definitions	Description and example
Universe	disco:Universe	**Universe** is the total membership or population of a defined class of people, objects or events. A population can be the target population (the

		universe that needs to be represented) and coverage or survey population (who are actually given an answer.)
Coverage	A subclass of Universe	A part of the Universe that is represented in our Study.
Agent	a class comprises people, either individually or in group, and entities who can perform actions Equivalent to Agent in RiC-O, DC, Schemaand PROV. Not equivalent to Actor in CIDOC!	Any human dead or alive, corporate bodies, organisations, software/machine used for an action *Iris Szeghy the composer.* *A photo is digitized using a specific machine*.
Group	Subclass of Agent Group of humans acting collectively Equivalent to Group in RiC, and CIDOC, Organization in Schema and PROV	corporate bodies, organisations, societies etc. *Greenpeace*. *Slovak Music Center*. *The Beatles.* *24VII EVO s.r.o.* a music publisher in Slovakia.

The Universe and the Coverage of the survey is a group of agents (humans or corporate bodies). When we survey orchestras, the universe is all the orchestras that conform to our universe definition (musical groups that have performed at least one concert in the timespan in Slovakia.) That is in itself a subgroup of those musical groups that were inactive in the timespan.

A.4.2 General model of agency



In the Open Music Registers we aim to register agents of music who can participate in activities, events, and create economically valuable works, recordings, print music, or live performances.

Agents, like persons, their groups and corporate bodies, create material entities, start processes, participate in events. Our model has a high-level interoperability because it models agency similar to rather foundational ontologies that are re-used in [the field of music] in the GLAM sector .

Class	Definitions	Description and example
Agent	a class comprises people, either individually or in group, and entities who can perform actions.	
Group	Subclass of Agent Group of humans acting collectively	The Slovak Youth Orchestra
Human	Subclass of **Agent** A class for any human being dead or alive, not including imaginary characters. Equivalent to Person in PROV, RiC, Schema, E21 in CIDOC	*Iris Szeghy the composer.* Celeste Buckingham the musician.
Corporate Body	Subclass of Group A group of humans who work together in the field of music.	Examples: Formal enterprise: 24VII EVO s.r.o.

	Many corporate bodies, for example, symphony orchestra, do not have a formal enterprise, but work, for example, within a broadcasting organisation.	Corporate body: The Slovak Youth Orchestra

Surveying corporate bodies that are incorporated is relatively easy; not-incorporated or informal corporate bodies is more difficult. Surveying agents that are humans sets different ethical and data protection limitations.

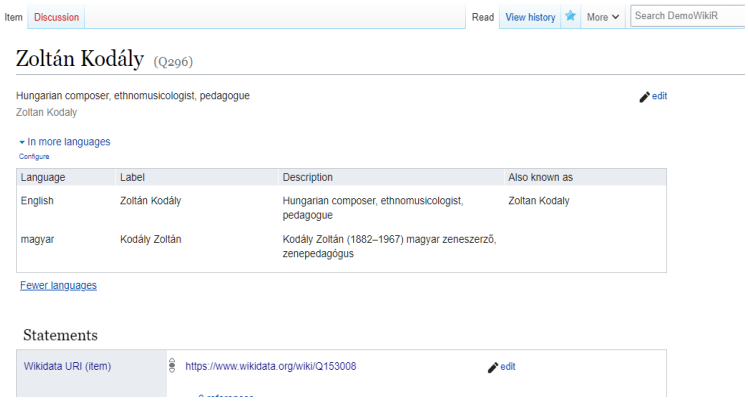
A.4.3 Human agents: musicians, music professional

Class	Property	Property definition	Example
Human	Given name Has equivalent in Schema and FOAF	Given or first name of the person	Wolfgang
Human	Family name Has equivalent in Schema and FOAF	Family or, last name of the person	Mozart
Human	Date of birth Has equivalent in Schema, FOAF, CIDOC and RiC	Date of birth of the person in YYYY-MM-DD format	W. A. Mozart was born 1756. 01. 27.
Human	Date of Death Has equivalent in Schema, CIDOC and RiC	Date of death of the person in YYYY-MM-DD format	W. A. Mozart died 1791. 12. 05.
Human	Place of Birth Has equivalent in Schema and RiC	Place of birth of the person	W. A. Mozart was born in Salzburg, Austria.
Human	Place of Death Has equivalent in Schema and RiC	Place of death of the person	W. A. Mozart died in Vienna.
Human	Gender Has equivalent in Schema and FOAF	Gender of the person at birth	man or woman W. A. Mozart was a man

One unusual element of a register that administers copyrights is that it contains data on deceased people and their heirs because the economic rights embedded in copyrights have a protection term beyond the author's death, and they are inherited. Consequently, a copyright register contains people who are neither composers, producers or performers of music but heirs without any music-related occupation. Some royalty payments leave the universe of living artists and go to heirs.

In event-based conceptual models, some of the properties of humans would be modelled differently. We use a simplified model that leaves out technical events in the database. For example, the **Birth of Ludwig van Beethoven** is an event, which had participants (mother, son), and an unknown place and time. While we do not know the time, we know that it must have taken place before the baptism of Beethoven (which is recorded with place and time.)

For rich inferences, we would export our database into a graph that assigns dates to events. For our purposes, we directly link W.A. Mozart or L.udwig van Beethoven to dates of birth, baptism, or death, but if we want to expand our knowledge about their vital circumstances, we may need to create events.



The agent data is described in several languages. It can be searched, validated, extended with the citizen-scientist level Wikibase graphical user interface, or with the MediaWiki API and SPARQL queries.

The general rules that we follow:

- Human Agent -> Birth (Activity) - > Timespan (date with day precision) + Location + Parents (if needed) is simplified to Human agent -> date of birth, place of birth.
- Corporate Body Agent -> Inception (Activity) -> Timespan (date with day, month, or year precision) + Location + Founders present, is simplified to corporate body founder, inception date.
- Human Agent -> Death (Event) -> Timespan + Location + Heirs (with respect to copyrights) is currently simplified to date of death.
- Corporate Agent -> Succession (Event) or Dissolution (Event) -> Timespan is simplified to successor and date of dissolution.

A very important aspect of human agents is their role or occupation in music. A role is a professional participation in music processes and activities, which, if habitual, becomes an occupation.

The *International Standard Classification of Occupations* (ISCO) is a four-level classification of occupation groups managed by the International Labour Organisation (ILO). Its structure follows a

grouping by education level. The two latest versions of ISCO are ISCO-88 (dating from 1988) and ISCO-08 (dating from 2008). ISCO is implemented in a largely harmonised way all over the world with relatively small differences. The ESCO (European Skills, Competences, Qualifications and Occupations) is the EU's definition (Commission 2017) , while the SSC (Standard Skill Classification) which is adapted specifically for the UK job market. (Elias, Dickerson, and Bachelor 2023).

A.4.4 Informal corporate bodies: orchestras, informal enterprises

The modern definitions used by the OECD, ILO, or IMF recognise the importance of the informal economy, and do not see it as an anomaly, but see the informal economy as a core component of economies at every level of development:

"the informal economy refers to all economic activities, excluding illicit activities, by workers and economic units that are, in law or in practice, not covered or insufficiently covered by formal arrangements […] While the informal economy exists everywhere, it is more prevalent in low-income countries, where it represents 89 per cent of total employment, compared to 82 per cent and 50 per cent, respectively, in lower-middle and upper-middle-income countries and 16 per cent in high-income countries." (International Labour Office 2015)

From a methodological point of view, we surveyed the problems and solutions of the high level of the music sector in (Antal, Kmety Barteková, and Remeňová 2023) as a methodological preparation of this work.

Our earlier research shows that in the 12 European countries, including Slovakia, a very high percentage of commercially released sound recordings are registered outside of the national domain of the ISRC registry.

In Slovakia, the majority of the cases fall outside the SK domain; registration outside of the country is widespread in Hungary and Poland, too, and it even occurs in the United Kingdom with independent labels that use foreign distributors. This is very logical: small member states like Malta, Estonia, or Latvia with a small population and music sector have very few entities capable of registration and commercialisation and artist working in such smaller countries tend to work with commercial agents from larger countries. (In the case of the United Kingdom, this larger country is the United States of America.) It is nearly impossible to say who has released a sound recording within a year and acted in the capacity of a music label in a country.

The best way to come to an almost comprehensive register is via the neighbouring rights societies, which have similar transparency requirements as author's societies, and they need to make public how they give them the mandate to collect neighbouring rights royalties from the broadcasting, transmission and various public performance locations. A music label with a serious business presence in a country such as Slovakia is expected to make an effort to collect such revenues, which, especially in emerging markets, maybe more than 50% of their total revenue (after adding private copying compensation.)

The high level of informality in both the live and the recorded sector means that many agents act as sole proprietors. In such cases, they may even use the same name they use in private life. The challenge is to protect their right to privacy as private persons and, at the same time, maintain transparency for their economic activities. A sole proprietor who releases music as a self-releasing or self-publishing author with a foreign distributor or publisher within the European Union is not only registered in the country's enterprise and tax register but also in the VIES standard VAT register of the European Union and Northern Ireland.

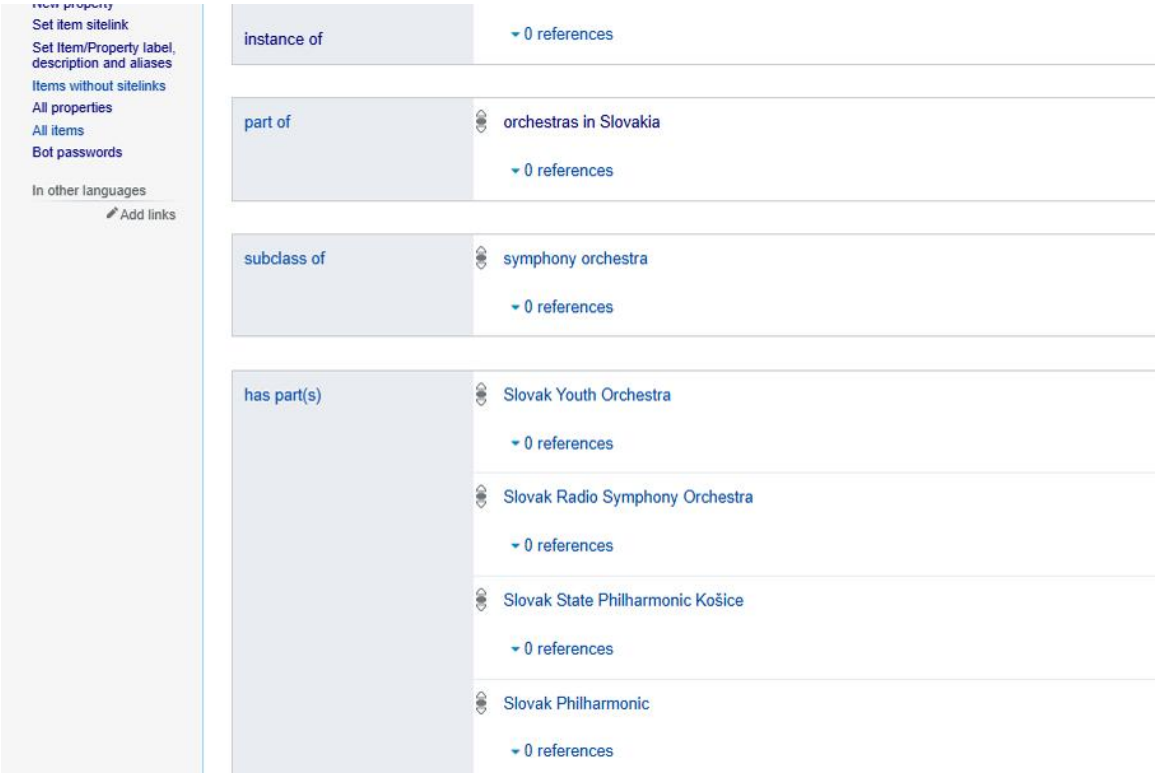
In such cases, we can cross-check with such state/EU registers, and identification is possible, but GDPR must be addressed, and the various interests of music business transparency, freedom of research and right to privacy must be balanced. Our register can identify these agents, but it must be addressed which part of this register can be made public and what uses are allowed. This question should be dealt with as the legal interoperability of our register.

Class	Property	Property definition	Example
Group	Inception Has equivalent in Schema and CIDOC	Time when an entity begins to exist	The Slovak Quartet (1956) was founded in 1956.
Group	Cease to exist Has equivalent in Schema and CIDOC	Time when an entity ceased to exist.	The Slovak Quartet (1956) disbanded in 1983.
Group	predecessor Has equivalent in RiC	Agent is predecessor of other Agent	The Slovak Quartet (1956) is the predecessor of the Slovak Quartet (2009)
Group	successor Has equivalent in RiC	Agent is successor of other Agent	The Slovak Quartet (2009) is the successor of the Slovak Quartet (1956)
Group	founded by Has equivalent in CIDOC	Group was founded by the Agent(s)	The Slovak Quartet (1956) was founded by Milan Telecký
Group	member Has equivalent in CIDOC	Agent is member of the Group, Group has member Agent	Jozef Horváth is a member of The Slovak Quartet (2009)

Groups are connected to humans. For example, to register an orchestra into VIAF, the register of national libraries, we must give an inception date and the founding members.

> *The Beatles* evolved in 1960 from John Lennon's previous group, the *Quarrymen* (predecessor), and built their reputation by playing in clubs in Liverpool and Hamburg, Germany, starting in 1960, initially with Stuart Sutcliffe playing bass. The core trio of Lennon, McCartney and Harrison, together since 1958, went through a succession of drummers, including Pete Best, before inviting Starr to join them in 1962.
>
> *The Beatles* has an inception in 1960 with four members: Lennon, McCartney, Harrison and Best. In 1962, before their first commercial release, Best is replaced by Starr. All studio recording releases define The Beatles as a group of Lennon, McCartney, Harrison, and Starr, however, after the dissolution of the band in 1969-1970, several recordings were released to the public that were recorded originally before 1962. Therefore, there are some known Beatles recordings where Lennon, McCartney, Harrison, and Best form the group.

Orchestras and bands that do not perform only in an amateur, educational or liturgical function must be able to receive payments for their performances. It is necessary that the musical group as a corporate body is linked to an entity that can issue invoice and account for the revenues and their taxes.



Musical groups (orchestras, choirs, bands, duos) are best modelled as informal corporate bodies, because their informal and formal structures often do not overlap. Example from a subclass of the Slovak musical groups, symphony orchestras at https://reprexbase.eu/demo/index.php?title=Item:Q1165>

In statistical terms, we are looking for various agents that produce sound recordings to make them available to the public. On mass broadcasting, retransmission and streaming platforms, a sound recording can only be made available if it is commercially released and has an ISRC identifier.

Such agents are sole proprietors, formal and informal enterprises (labels), often working within a larger corporate entity, and sometimes, they are incorporated enterprises with a relatively homogeneous economic activity profile.

A.4.5 Musicians works and recordings

The purpose of our surveying register is to help to collect data on the agents of music. We develop our register in a way that human agents in composer, producer and performing roles can be connected to their works, recorded, print and live public performance manifestations. We will show our work and manifestation register in a separate document.

> **Is A Work in the Public Domain?**
>
> A possible application of SKCMDb is to find out if a work needs a license or it is in the public domain.
>
> The query needed for this is connecting a Musical Work (Symphony No. 9.) to all known composers (Ludwig van Beethoven) and to all date of death properties. If all of them are earlier than the calendar year number for the previous year minus 70, i.e., 2024-70, we conclude that the work is in the public domain.
>
> For the purpose of the economy register, a work register is not necessary, but for most uses of the SKCMDb, this connection is needed.

A.4.6 Formal corporate bodies: municipalities, enterprises, foundations

Concert promoters are usually easier to observe than orchestras, but their form varies. They can be enterprises with an economic activity (classified with a NACE code), NGOs, or public bodies, such as municipalities. Concert promoters are found in several state registers, and crucially, they should be found in the licensing database of author's societies. Only those concert promoters who put on stage out-of-copyright or exclusively liturgical music (in churches, synagogues, etc.) are exempted.



A formal corporation from the music publishers' register.
<https://reprexbase.eu/demo/index.php?title=Item:Q201>

The concert producers (as orchestras) should be somewhere in the register of enterprises or sole proprietors, but finding them is looking for a needle in the hay. They may use a company of one of the members for this purpose, which has an entirely independent main economic activity from R90 performing arts. We will show the difficulties of capturing live music events with the hypothetical example of Amber Smith, a band, playing in a free show of organised by Nagymaros, a municipality. The members of Amber Smith do get paid.

> **Concert economics: supply side**
>
> There is a band called Amber Smith (band). It is a corporate body (Group) in CIDOC-CRM, an agent in DCTERMS.
>
> The band has a corporate body, Ambersmith Kft. that is a legal entity. It is owned by the members of Ambersmith (band).
>
> The Ambersmith Kft. has main economic activity of Performing Arts.

```
# Ambersmith (band)
<http://example.org/ambersmith-band> a crm:E74_Group ;  # Corporate body (group)
  dcterms:type <http://purl.org/dc/dcmitype/Agent> ; # Agent in DCTERMS
  schema:name "Amber Smith" .

# Ambersmith Kft. (legal entity)
<http://example.org/ambersmith-kft> a crm:E74_Group ; # Corporate body (group)
  crm:P127_has_owner <http://example.org/ambersmith-band> ; # Owned by the band
  schema:name "Ambersmith Kft." ; # Optional name
  crm:P19_was_intended_use_of <http://example.org/nace-r900> . # Main economic activity

# NACE code for Performing Arts
<http://example.org/nace-r900> a skos:Concept ; # Using skos:Concept for NACE code
  skos:notation "R900" ;  # The NACE code
  skos:prefLabel "Performing Arts"@en . # Label (optional)
```

If we want to measure concert economics, we need to be able to see the payment transaction connected to the payer or the payee.

Who may have a record with both agents (payer and payee?). The author's rights organisations must license this public music performance so they have information about both the payee and the payer.

Concert economics: demand side

There is a corporate body, the municipality of Nagymaros.

Nagymaros organises an Event called “Free Concert”.

Nagymaros pays for the performance of Ambersmith (band) to Ambersmith Kft.

A potential complication is that actually a non-profit company owned by the municipality, Nagymaros Cultural Events Non-Profit Kft pays the Ambersmith Kft..

```
# Municipality of Nagymaros
<http://example.org/nagymaros-municipality> a crm:E74_Group ; # Corporate body
  schema:name "Municipality of Nagymaros" . # Optional name

# "Free Concert" Event
<http://example.org/free-concert-event> a crm:E5_Event ; # Performing Arts Event
  crm:P9_consists_of <http://example.org/nace-r900> ; # Classify as Performing Arts (optional)
  crm:P7_occurred_at <http://example.org/nagymaros-municipality> ; # Occurred at Nagymaros
  crm:P11_had_participant <http://example.org/ambersmith-band> . # Ambersmith (band) participa
ted


# Nagymaros initiated the event
<http://example.org/nagymaros-municipality> crm:P13_has_occurrence <http://example.org/free-c
oncert-event> .  # Nagymaros has the event
```

The connection between the live music production (Amber Smith and Ambersmith Kft) and the staging of the music (Nagymaros municipality and its non-profit cultural entity) are connected in space and time if Artisjus, the Hungarian collective rights management agency, licenses. In this case, the licensing database contains both important Agents, the time, space, and likely some other variables like maximum capacity of the location, if the event is ticketed or not, if ticketed, and what the ticket price is.

Concerts are very often organised by entities that are not for-profit businesses and do not have a NACE code. Non-profits use the *International Classification of Non-Profit Organizations (ICNPO)*, and the European Union uses various classification for pubic sector entities such as the municipality of the town of Nagymaros.

The Statistical Classification of Economic Activities in the European Community, commonly referred to as NACE (for the French term “nomenclature statistique des activités économiques dans la Communauté européenne”). It is derived from the UN classification ISIC, revision 4.

The Statistical Classification of Economic Activities in the European Community, commonly referred to as NACE (for the French term “*nomenclature statistique des activités économiques dans la Communauté européenne*”), is the industry standard classification system used in the European Union. The United Kingdom has a slightly different version of NACE called

At the time of Brexit, the EU (and the UK) used the second revision of NACE (Eurostat 2008), currently, the EU prepares for the rollout of the the revision 2.1 in 2025 (ESTAT 2022), while the UK uses the United Kingdom Standard Industrial Classification of economic activities (UK SIC) (National Statistics, n.d.).

Just like with the occupation taxonomies, the economic activity taxonomies are hierarchical, so the EU, the UK and other regions will be aligned by ISIC to a certain point. The EU itself is aiming alignment at the level of four-digit precision; and member states (and the UK) often create five- or six-digit precision versions.

One problem with applying NACE is that it is designed for larger corporations and does not work very well with microenterprises that show some level of informality (and does not apply to informal enterprises.) Microenterprises often have an unclear economic activity profile, and their data is also lacking in detail. A larger company that applies the IFRS financial reporting standards (or its UK or any European national form) must report specific segments of its business separately. However, a company smaller than the "small" enterprise threshold falls under a simplified accounting (and tax reporting) scheme in perhaps every European country. Some countries must only report a net turnover and profit after tax. In the United Kingdom, four scenarios may apply to microenterprises, and in one case, it must not even report any flow variables (like net turnover or profit after tax.) (Chartered Accountants in England and Wales 2023)

The use of NACE is particularly problematic for music because it does not provide a clear mapping for music corporate bodies that fall under the enterprise definition. (The same problem applies to some music occupations), and even if it does, it will identify some data poor subjects. A bigger problem is that many music corporate bodies do not have a NACE code, for example, because they work as an "informal enterprise" or as a non-profit entity or a public body. NACE appears to be somewhat suitable to identify music publisher, but hardly suitable for music labels, and perhaps entire unsuitable for concert promoters and concert venues.

A.4.7 Business-to-business users of music

NACE is not absolutely useless, of course. Identifying some licensed user groups of music, such as radios, is rather good. Sole proprieators cannot operate radios in most member states, and the market share of community radios operated by small non-profits (who do not have a NACE code) is relatively low. The bulk of radio stations that pay a licensing fee for broadcasting music usually have a NACE code that is well mapped to their activities. The same can be said about television broadcasters and cable retransmission entities.

Often, the research question in business or academia is not related to music creation, but the values the creations in their business use. This leads to registering a different set of agents, i.e., agents who use music in business-to-business relations. In cultural access and participation surveys, we also may ask natural persons as music users.

Corporate user agents have a simple typology in the use of the mechanical licensing mode; these are producers of physical records or labels who benefit from the collective management of authors' rights for mechanical copying.

A massive group of music enterprises are the agents that use music in public performance and retransmission.

- ☒ Public performance of live music events (foreground music): festival organisers and permanent venues that stage music; either with a main focus on music staging (concert halls, music clubs) or multi-purpose cultural or entertainment venues such as arenas and cultural establishments. Such agents are usually part of the formal economy, but they are often not corporations but municipalities, associations, foundations, and religious associations. Informality is an issue.

- ☒ Broadcasting and retransmission of recorded music via radio, television, cable and satellite television, in which cases we deal with a relatively low number of well-established users who may have their registered, for example, because national authorities register and distribute access to the broadcasting wave spectrum. Informality should not be an issue; private persons very rarely operate private radios or TVs.
- ☒ Public performance (background music) of recording music, such as restaurants, hotels, retail outlets, various public buildings like malls, health facilities, workplaces like warehouses. Informality is an issue in the restaurant subsector.

A.5 Experimental design, materials and methods

ⓘ

Our connected registers are, to our knowledge, unique; they utilise recently novel statistical production and legal regulatory innovations to connect public and private registers. Our database's data governance and exchange are unique because of the serious data protection restrictions posed by the use of both public statistical and private registers.

A *statistical register* is a continuously or regularly updated set of objects for a given population. It contains information on the identification and accessibility of population units as well as other attributes which support the surveying process of the population. It serves as a constantly updated list of potential data sources: people or enterprises, for example, who may be invited to a sample survey or a census. The statistical register is both a coordination tool for data collection (everybody is found who should provide data), and an important data quality management tool (we know if somebody was not found, how it will distort our resulting datasets.) For example, as earlier stated, Statistics Norway applies about 100 statistical registers.

Our primary concern in Slovakia is the creation of a music business register, because this could provide indicators for the public policy-level and institutional/enterprise level implementation of the Slovak cultural strategy.

The authoritative source on statistical business registers is the *United Nations Guidelines on Statistical Business Registers. Final draft prior to official editing* (United Nations Statistics Division 2020) which is heavily based on the former UNECE *Guidelines on statistical business registers* (UNECE 2015). The European guideline is the *European business statistics methodological manual for statistical business registers. 2021 edition* (Eurostat 2021b).

Our register is designed to work with relational databases that are regularly used in rights management or GLAM organisations to manage membership lists and catalogues and, at the same time, work with graph databases that are increasingly replacing relational databases for supporting trustworthy, human-controlled AI.

As a conceptual model, we use the Wikibase Data model, which is described on the ontological level. For the sake of interoperability, we created classes and relationship types that follow strict equivalence with the GLAM standard ontologies and, whenever possible, with Polifonia Music Match and the Music Ontology.

For music economic research, further extensions are needed to the GSIM statistical information model (for surveying and processing answers, see Chapter 3) and the ISIC-ESCO and ISIC-NACE-SIC classification systems.

A.5.1 Occupations

The European Union recently published its ESCO taxonomy in an ontological form, (De Smedt 2023), which is already connected to two registers (the *Awarding Body* register for those European institutions that can award a qualification and *Work Context Awarding Bodies*.) Awarding bodies are not very important in music, because most musicians acquire skills by informal learning practices (Green 2017).

The ISCO/ESCO systems have the following pillars:

- Occupation (these relate to habitual roles)
- Skill (and competences)
- Qualification

We utilise the ESCO ontological form, because it can support any ISCO-originated taxonomy with a clear interoperability with registers. The ISCO (and its ESCO and SOC) implementations as taxonomies are hierarchical: on the top level, every UN member state applies them the same. Beyond a certain level of granularity, the EU and UK can diverge, or new applications can be built. We actually aim to build a divergent taxonomy that adds further granularity to the point where the EU and UK definitions are still aligned.

When we want to understand the relative economic, environmental or social weight of music activities and processes, we must be able to use a systematic taxonomy of such occupations and roles. The current international occupation taxonomy, ISCO and its European ESCO extension are not sufficiently detailed for understanding music.

In the creation of new musical work, we are talking about natural persons who have an occupation as a *composer*. In ESCO 2652.1 describes the *composer*, and 2652.1.1 describes the *music arranger*, and 2641.4.1 the *lyricist*. (Most composer do not make a living out of composing only, and this is not their only occupation; so one human agent usually has several parallel roles or occupations.)

Lyricists interpret the style of a music piece and write words to accompany the melody. They work together with the music composer.

A note: in the current ESCO definition the composer *excludes sound artist, lyricist and music arranger.*; the exclusion of the lyricists is ambiguous. Lyrics can be registers as a literary work and as part of a musical work. It needs further validation if this ESCO taxonomical difference causes problems in our data model.

In ESCO, 2654.3.1 *incorrectly* describes a music producer as

Music producers are responsible for acquiring music to be published. They listen to demos of songs and determine whether they are good enough to be published. Music producers oversee the production of records. They manage the technical aspects of recording and editing.

In fact, music producers deal with the release of recordings, and not with the publishing of musical works.

As human agents, in the ESCO definition: “musicians perform a vocal or musical part that can be recorded or played for an audience. They have know-how and practice of one or many instruments or using their voice. The musician can also write and transcribe music.”

So, if we want to measure consistently the economic activities of music professionals (as natural persons agents), we should come up with a description of habitual occupations or roles within these occupations that are understood the same way in the United Kingdom or Slovakia. This is beyond the scope of this document, but in the background of Open Music Europe there is plenty of experience for doing this for the purpose of harmonised surveying of music professionals.

Our approach to understanding a music professionals occupation is a two-stage filtering, which starts with the Eurobarometer standardised, non-country-specific general occupation question.

Eurobarometer-equivalent general occupations

English	Lithuanian
What is your current occupation? Which statement best describes your current employment situation?	Kokia jūsų dabartinė profesija? Kuris teiginys geriausiai apibūdina dabartinę jūsų darbo situaciją?
1a - Musician, composer, songwriter, DJ … can be selected in more detail on Question 4!	1a - muzikantas, kompozitorius, DJ… daugiau detalių ketvirtame klausime
1b - Other independent artistic activities, choreographers, graphic artists, etc.	1b - kiti nepriklausomi artistiniai užsiėmimai - choreografija, grafikos dizainas ir t.t.
1c - Music technician, eg sound engineer, light technician, supplier … clarifies the next question Q4	1c - muzikos technikas, pvz. garso inžinierius, apšvietėjas, tiekėjas ir t.t.
1d - Music-related entrepreneur (label, publisher, booking, etc.) … clarifies the next questions Q4	1d - su muzika susijęs verslas (leidėjas, booking agentas, įrašų kompanija)
1e - Employed manager of a music-related company	1e - su muzika susijusios kompanijos vadybininkas
1f - A non-managerial employee of a music company	1f - su muzika susijusios kompanijos “ne-vadybinis” darbuotojas
1g - Full-time student in music or music related education	1g- Muzikos arba su ja susijusio dalyko studentas
2 - DOES NOT WORK in general, the more detailed the following	2 - nedirba, daugiau detalių toliau
2a - Responsible for ordinary shopping and looking after the home, or without any current occupation, not working	2a - atsakingas už apsipirkimą ir rūpinimąsi namais arba šiuo metu neturintis darbo
2b - Student	2b - studentas
2c -Unemployed or temporarily not working	2c - bedarbis arba laikinai nedirbantis
2d - Retired or unable to work through illness	2d - pensininkas arba negalintis dirbti dėl ligos

3 - SELF EMPLOYED in general, more detailed below	3 - dirbantis savo paties įmonėje, daugiau detalių toliau
3a - Farmer, fisherman	3a - ūkininkas, žvejys
3b - Highly qualified professional (lawyer, medical practitioner, accountant, architect, etc.)	3b - aukštos kvalifikacijos profesionalas (teisininkas, medikas, architektas ir t.t.)
3c - Owner of a shop, craftsmen, other self-employed person	3c - parduotuvės savininkas, amatininkas arba kitas laisvai samdomas asmuo
3d - Business proprietors, owner (full or partner) of a company	3d - verslo ar kompanijos savininkas
4 - EMPLOYED in general, more detailed below	4 - dirbantis, daugiau detalių toliau
4a - Employed professional (employed doctor, lawyer, accountant, architect or similar)	4a - įsidarbinęs profesionalas (įsidarbinęs daktaras, teisininkas, architektas ar pan.)
4b - General management, director or top management managing directors, director general, other director)	4b - generalinis direktorius arba kitas direktorius
4c -Middle management, other management (department head, junior manager, teacher, technician)	4c - vidurinės arba kitos grandies vadybininkas (departamento vadovas, jaunesnysis vadybininkas, mokytojas, technikas)
4d -Office Worker	4d - ofiso darbuotojas
4e - Employed position, not at a desk but travelling (salesmen, driver, etc.)	4e - darbas keliaujant (pardavėjas, vairuotojas ir t.t.)
4f -Service, public service (eg hospital, restaurant, police, fire department)	4f - darbas viešų arba privačių paslaugų sektoriuje (ligoninė, gaisrinė, restoranas, policija ir t.t.)
4g - Group leader, foreman	4g - grupės lyderis
4h - Specialist	4h - specialistas
4i -Physical worker, home staff, baby sitter	4i - fizinis darbas, vaikų auklė
5 - NEVER DID PAID JOB, and not student, not retired	5 - niekada nedirbo apmokamo darbo, ne studentas, ne pensininkas

As a second step we establish artistic roles which can be seen as extensions of the ISCO/ESCO 235 class. More formal mapping and extension could be carried out, but given that labour force statistics are usually not available with more granularity, the current level of mapping is sufficient for comparison with official statistics.

As an example, we show the CEEMID survey codebook in English and Lithuanian. The codebook is also available in Bulgarian, Croatian, Czech, Estonian, German, English, Hungarian, Polish (needs editing), Romanian, Serbian, and Slovak.

The following grouping works are harmonised ISCO, European ESCO, North American or UK occupation taxonomies, too. This means that a survey that is conducted among music professionals should result compatible results with labour force statistics in the same country at least on the granularity ISCO/ESCO 2 or 3 digit levels.

ESCO 235 creative and performer occupations

musician and DJ roles	**muzikantų ir DJ rolės**
11-L - performing musician, not author	11-L pasirodymų muzikantas, ne autorius
12-L - performing DJ, no own music	12-L - pasirodymų DJ, jokios savo muzikos
13-C - composer	13-C kompozitorius
14-R - producer	14-R prodiuseris
15-L - other artistic role	15-L kita artistinė rolė
16-C - lyricist	16-C žodžių autorius
17-L - author and composer	17-L autorius ir kompozitorius
18-L - DJ producer with own music	18-L DJ prodiuseris, su savo muzika
19-L - conductor	19-L dirigentas

The following roles mainly provide more precise descriptions to ESCO 3521 occupations.

Technician roles, relate to ESCO 3521 subgroups

technical role	**techninės rolės**
31-L - sound engineer and team (live music)	31-L garso inžinierius ir komanda (gyva muzika)
32-R - sound engineer and team (studio)	32-R garso inžinierius ir komanda (studija)
33-L - light engineer and team	33-L šviesų inžinierius ir komanda
34-L - stage master and team	34-L scenos prižiūrėtojas ir komanda
35-L - other technical (live)	35-L kita (gyva muzika)
36-R - other technical (studio)	36-R kita (studija)

The following codes relate to ESCO 12 managerial roles. They could be mapped more tightly to ISCO/ESCO.

Managerial roles, relate to ESCO 12 roles

music management roles	**muzikos vadybos rolės**
41-L - tour manager	41-L turo vadybininkas
42-L - concert promoter in an institution	42-L koncertų reklamuotojas institucijoje
43-L - independent concert of festival promoter	43-L nepriklausomų koncertų ar festivalių reklamuotojas

44-L - booking agent	44-L užsakymų (booking) agentas
45-C - work in a collective management society	45-C darbas gretutinių teisių asociacijoje
46-R - work in a music publisher	46-R darbas muzikos leidyboje
47-R - work at a record label	47-R darbas įrašų leidybos firmoje
48-L - work at artist/talent manager	48-L darbas atlikėjų vadybininku
49-L - work at music club or discotheque	49-L darbas muzikos klube ar diskotekoje
50-L - other management roles (live music)	50-L kitos vadybinės pareigos (gyva muzika)
51-R - other management roles (recordings)	51-R kitos vadybinės pareigos (įrašai)
52-C - other management roles (compositions)	52-C kitos vadybinės pareigos (kompozicija)
53-L - transporter for bands, orchestras	53-L grupių/orkestrų vairuotojas
61-L - other workers (music events)	61-L kitas darbuotojas (muzikos renginiai)
62-R - other workers (recordings)	62-R kitas darbuotojas (įrašai)
63-C - other workers (compositions)	63-C kitas darbuotojas (kompozicija)
64-O - other artist roles in live performances	64-O kita rolė gyvuose pasirodymuose

Then we have the music educator roles which relate to ESCO 2354, and some other roles.

Various other roles in music education (ESCO 2354) and other

music education and journalism roles	**muzikos edukacijos bei žurnalistinės pareigos**
80-E - music journalist	80-E muzikos žurnalistas
81-E - music educator	81-E muzikos mokytojas
82-E - work in a rehearsal studio	82-E darbas repeticijų studijoje

In our surveys, we always ask for a primary and secondary role, and the secondary roles often relate to the broader music ecosystem.

Used for typical secondary roles

other roles and student	**kitos rolės arba studentai**
91-M - work in a record shop	91-M dirbu įrašų parduotuvėje
92-M - work in an instrument shop or repair shop	92-M dirbu instrumentų arba jų taisymo parduotuvėje
98-M - full time student	98-M studentas
NA - not applicable or none above	NA - netinka nė vienas iš variantų

For compliance with the UK Music Surveys we used the following level of competenece questions:

Self-assessed level of competence

Would you describe yourself in this role…	Ar jūs apibūdintumėte save…
1 - amateur	1 - mėgėjas
2 - learning to be professional	2 - mokausi būti profesionalu
3 -part-time professional	3 - profesionalas ne visą laiką (part-time)
4 - full-time professional	4 - profesionalas visą laiką (full-time)
NA - cannot or prefer not to say	NA - nežinau arba renkuosi neatsakyti

A.5.2 Economic activities of corporations

Another important classification is ISIC (UN) or NACE (EU) or SIC (UK). This is a taxonomy for the economic activities of enterprises. Some member states, like the Netherlands or Hungary, apply a national extension of NACE to sole proprietors, which overlaps with the ISCO/ESCO/SOC defined occupational activities, but only in the case of sole proprietors. For example, in the relevant taxonomy for Hungarian sole proprietors, those sole proprietors of music whose main economic activity is have a code of 900102, derived from NACE 9000. ISIC, NACE, or SIC is only suitable for the classification of some enterprises.

A.6 Data Curation Policy for Slovak Music Economy Register

1. Introduction The creators of Slovak Music Economy Register are committed to maintaining high-quality, accurate, and reliable data for research and analysis. This policy outlines the principles and processes governing the curation, management, and dissemination of data within the database.

2. Data Scope and Coverage

The database curates and maintains data relevant to the economic activities of musicians, composers, their corporate bodies, and music businesses operating in the territory of the Slovak Republic. We include agents who work as music performers or composers and who were either born in the Slovak Republic, use Slovak-language lyrics, or identify with Slovak or minority ethnic groups from the Slovak Republic — even if they are active abroad. For natural persons, inclusion in the register is subject to compliance with data protection regulations, including the General Data Protection Regulation (GDPR), which governs how personal data may be stored, processed, and shared.

The dataset consists of observations of music agents (rows) and their attributes (columns). Categorical information, such as the classification of the agent's activity or role, is organised using a modified version of the classification system developed by the Slovak Music Centre. This system has been collaboratively adapted by Anna Žilková, Anna Márta Mester, and Dániel Antal.

SKCMDb — the Slovak Comprehensive Music Database — is the technical platform used to implement the Slovak Music Economy Register. It is built on a MariaDB relational backend extended with graph capabilities via Wikibase, the same open-source platform that powers Wikidata. Wikidata, the world's largest open graph database, has been successfully used in the past decade to create interoperable library services (van Veen 2019; Bianchini, Bargioni, and Pellizzari di San Girolamo 2021; Sardo and Bianchini 2022), museum services (Pfeiffer and Gayo 2021), and archives (references), and creating interoperability among the institutional silos of these three types of heritage collections (Alexiev et al. 2020; Rossenova, Duchesne, and Blümel 2022). EU institutions themselves also use it (Fischer, Thill, and Angjeli 2020).

A.6.1 Data Collection and Ingestion

A.6.1.1 Primary data sources

The population register is the primary data source for labour surveys and other personal surveys in the Slovak Republic. Researchers cannot access this register directly.

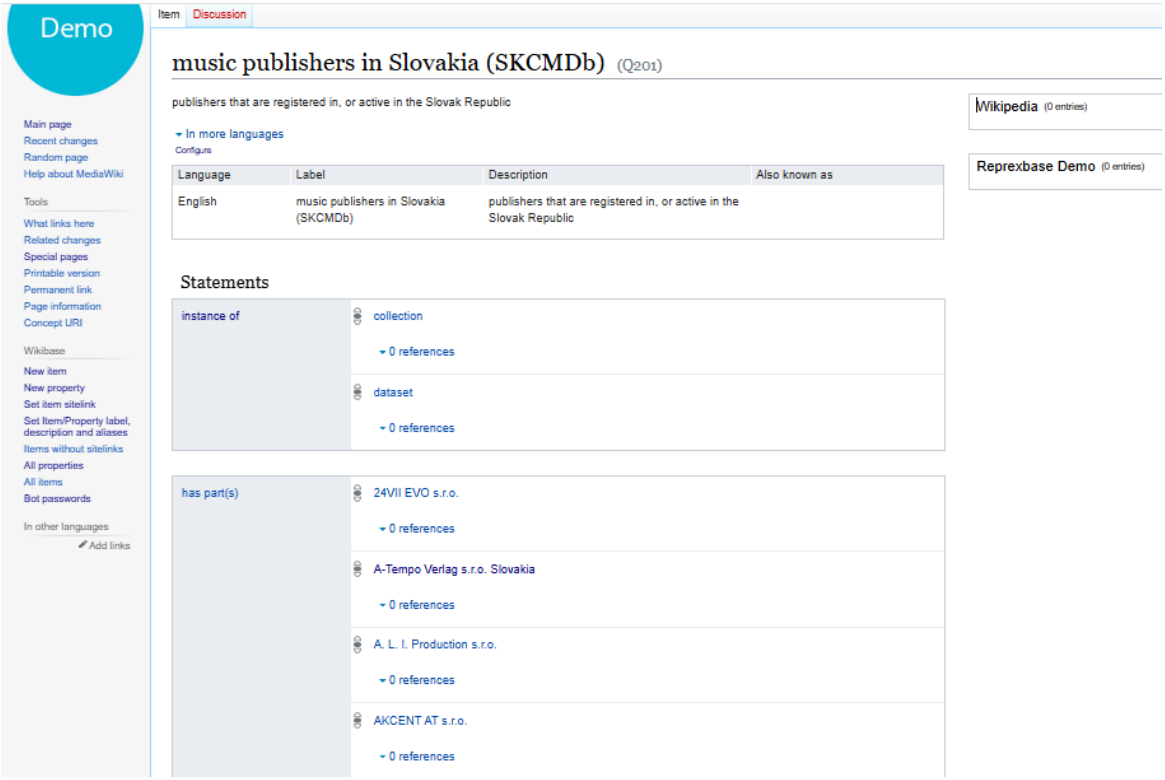
For **Human** agents, we use

- ☐ the database of SOZA contains all living and deceased persons known to hold copyrights in Slovakia, including personal identifiers, rights metadata, and registration histories, stored in relational database tables;
- ☒ the database of the Slovak Music Centre, which contains information about living and deceased human persons as composers and performers (including music directors or conductors) and educators in relational database tables;
- ☒ Wikipedia contains data about notable human agents in a loosely structured form of narrative encyclopaedic entries,
- ☒ Wikidata contains data about notable human agents in graph format.
- ☐ The database(s) of sole proprietors who are registered to conduct regular business activities and may be registered for European community VAT transactions. We are still investigating the legal implications of using this data source.

For **Corporate bodies**:

- ☒ The registered music publishers are in the SOZA database.
- ☒ The database of the Slovak Music Centre, which contains information on certain musical groups
- ☒ The registration data of IFPI Czechia and Slovakia.
- ☒ The company house register of the Slovak Republic.
- ☒ The statistical register of the KULT surveys in the Slovak Republic for orchestras, record labels, and concert venue operators.
- ☒ Open Corporates is a global register of legal entities that contains all Slovak music businesses, and almost all European music businesses.

© 2023 OpenMusE | HORIZON-CL2-2022-HERITAGE-01-05 | Grant Agreement No. 101095295



The register of music publishers in Slovakia

For **Events**, we are are only experimenting with the connection of the following data sources:

- ☒ The concert calendar of the Slovak Music Centre. Historical data on events from the Slovak Music Centre.
- ☐ The non-public licensing database of SOZA that contains information about licensed music events.

For **Places**, we are only experimenting with the connection of concert venues with actual events, and participating music performers and concert promoters.

- ☒ The concert calendar of the Slovak Music Centre for locations. Historical location data on events from the Slovak Music Centre.
- ☐ The non-public licensing database of SOZA that contains information about the location of public performance venues.

A.6.2 Limitations

ⓘ

Our register contains data about living and deceased natural persons and informal and formal corporate bodies. The legal requirements to use the data vary; the use of human data subjects is governed by the General Data Protection Regulation of the European Union, among other regulations.

The SOZA membership register is not directly accessible. However, it is accessible to send out a survey to the people registered on this list no more than once a year. The statistical properties of the registered humans (such as the number of registered living persons, mean or average age, and distribution of certain income) are available for testing survey coverage. For example, it is possible to

test how closely anonymous responses in a survey correlate with the actual target population (universe), i.e., to test the representative coverage within the anonymised survey responses.

A.6.3 Data Quality and Integrity

- Data undergoes regular validation and integrity checks to ensure accuracy. This includes human validation by Anna Žilková as the head of the Department of Documentation and Informatics at the Slovak Music Centre, and integrity checks by the data stewards.
- Version control is implemented to track changes and updates; each published new version received a version-DOI on Zenodo.
- Erroneous or obsolete data is flagged, corrected, or archived appropriately.

A.6.4 Data Ontologies, Taxonomies, and Enrichment

Importantly, ESCO (European Skills, Competences, Qualifications and Occupations) and SDMX (Statistical Data and Metadata eXchange) are available in fully ontological forms (European Commission 2020a; SDMX Technical Working Group 2021). This means they are encoded in a formal, machine-readable structure that includes logic-based reasoning rules — allowing computers to automatically infer relationships and classifications. These ontologies support dynamic alignment, semantic validation, and intelligent linking across datasets.

In contrast, NACE (the EU's economic activity classification), the UK's SIC (UK Office for National Statistics 2022) and ISCO (the International Standard Classification of Occupations) are available as linked open taxonomies (Eurostat 2024; International Labour Organization 2024). While also machine-readable and publicly accessible, they lack the deeper logical structure of ontologies. They are essentially structured vocabularies — excellent for standardising terms and enabling cross-referencing, but without the capacity for automated reasoning or rich semantic querying.

Together, these resources enable the Open Music Observatory to embed music industry and cultural data in a broader ecosystem of interoperable, standards-aligned, and future-proof knowledge infrastructures — bridging national data with European and global statistical frameworks.

We use elements and patterns from five formal, general-purpose ontologies expressed in the standard ontology languages OWL, RDF (with extension of RDFS), and we also aim for interoperability with application of the Music Ontology and the Polifonia Music Meta.

Our business users, including music artists and their corporate bodies and enterprise will most likely need interoperability with Schema.org, which standardises most commercial web applications. Our memory institution partners will likely want interoperability with Dublin Core, *Records in Context*, or *CIDOC*; we believe that this is also rather suitable for most rights management tasks.

Our research users want to be compatible with GSIM, SDMX, and DDI, to create questionnaires, administer surveys that get responses to the questions, model the question representations to receive output variables and process them statistically to be used in various academic, business, or policy research.

The interoperability with such statistical applications is tangential but important: it is important that target population is well-described and the responses are well-connected to important attributes of the data subjects, such as their music occupations or economic activities in music.

Metadata schemas and controlled vocabularies are used to enhance data organization and discoverability.

© 2023 OpenMusE | HORIZON-CL2-2022-HERITAGE-01-05 | Grant Agreement No. 101095295

- Our primary interoperability target is museum applications, so we use the CIDOC-CRM ontology (Bekiari et al. 2024), abbreviated in this document with the crm prefix.
- Our secondary interoperability target is with library applications for cross-referencing research and data that cites the objects in the collections. We use the Dublin Core library ontology, DCTERMS for this purpose (Core 2020), abbreviated with the dcterms prefix in this document. DCTERMS provides a wide interoperability with data repositories, as they often use DCTERMS as their foundation ontology.
- Our third interoperability target is DataCite, which is the preferred ontology for European open science repositories. (Group 2021)
- Records in Contexts is the new archive standard (Archives Expert Group on Archival Description 2023).

For ensuring that researchers and their software applications understand our descriptions the same way the descriptions of garments and their provenance information, we use the following controlled vocabularies.

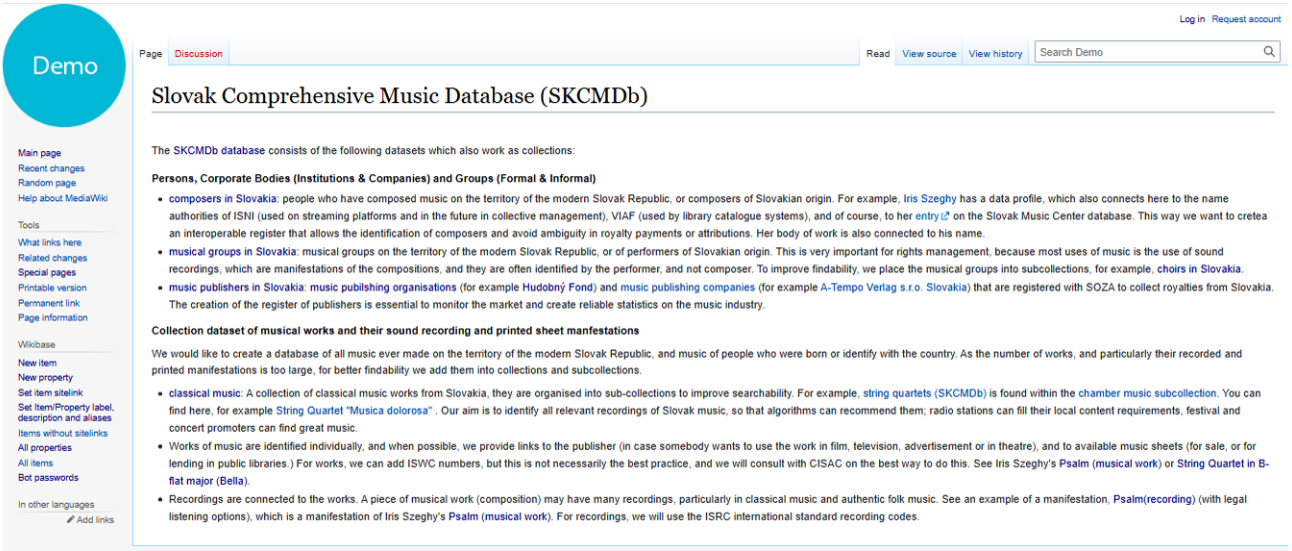
- For describing places (of original provenance and current hold locations), we use Geonames, a global taxonomy of place names.

A.6.5 Data Updates and Maintenance

- The database is updated on a regular basis, with at least one update in every quarter (three months.) We may update more frequently if we acquire significant new data.
- Changes in data sources or methodologies are documented transparently.
- Regular audits are conducted to identify and resolve inconsistencies.

For manual updates, see our Working with Persons in Wikibase tutorial.

A.6.6 Data Access and Sharing



Temporary opening page here. We will change this to the final (not demo) page by the end of February.

We define several levels of data access. For reading the data:

- ☒ Reading public datasets manually does not require password authentication. The data will be available with a Creative Commons license CC BY-NC-SA 4.0. We will provide a commercial policy, too, in the next update of the Dissemination, Exploitation and Communication deliverable.
- ☒ API access to mass-download data requires a "bot" username and password with a CSFR-authenticated session. This is technically possible, and we have rather detailed technical documentation, but we have not yet opened a request for these API users.
- ☒ SPARQL queries: no password is needed for public datasets (this function does not work yet; we plan this functionality to work by the end of March.)
- ☒ Read access to restricted data: requires an explicit agreement from data owners.

For writing the data:

- ☒ Suggesting manual corrections and extensions requires a username and a password to provide data provenance to data changes. Experimentally, we have provided manual user names and passwords and are working on the data stewardship model to review changes.
- ☒ Mass upload data via API: technically possible with a bot user, password, and CSFR token. Currently, only Reprex has mass upload rights.

A.6.7 Data Security and Privacy

- We do not handle personal data, with the exception of adding names of data curators, researchers, who must be attributed as creators of the dataset.
- Access to our datasets is not strictly controlled, however, only the dataset placed in the Zenodo scientific repository is authentic for re-use. We use Zenodo to create an independent, trusted, authoritative copy of the dataset with versions, identified by a new DOI for each new version.

A.6.8 Governance and Responsibilities

We follow the dataspace concept in building Slovak Music Economy Register. A dataspace is an emerging approach to data management which recognises that in large-scale integration scenarios involving many partners, it would be prohibitively expensive and time-consuming to obtain an upfront unifying schema across all sources or to come to a legal agreement on the terms of using or exchanging the data. It is an intelligent application that allows a near-instantaneous exchange, processing, sharing and provision of data on an "as-needed" or "as-permitted" basis while retaining complete control of each data holder over the conditions (e.g., who, when, and under what condition) of access to their data (Curry 2020; EBU and Gaia-X 2022, p16; Nagel and Lycklama 2021).

To translate this to laymen terms, it is unlikely that Slovak Music Economy Register will be allowed to have a permanent, live connection to national libraries, or the inventory systems of various archives and museums. But we hope that we can negotiate one-time or period, annual exchanges, supervised by both an interdisciplinary research team and data stewards. Applying the dataspace design allows that such one-time or regular synchronisation opportunities are well-planned, have no significant additional costs, and can be easily overseen by legal, data security and other experts. We build Slovak Music Economy Register as a system that can, when permitted, receive larger data tables from various museum, library, and archive systems, or open science repositories.

Dataspaces themselves can be connected with each other. Such connections are called data federations. Slovak Music Economy Register , by design, aims to be federated with Wikidata (the largest open knowledge graph database of the world), the Europeana Data Space (that connects various digital heritage databases in Europe), and the European Cultural Heritage Cloud, which will be an extension of the European Open Science Cloud (EOSC) to various digitised heritage collections (Directorate-General for Research and Innovation (European Commission) 2022). Data federations will allow us to import any new data from the Europeana Data Space into Slovak Music Economy Register as soon as it appears there, and the music experts or Slovak Music Economy Register will agree that the data is relevant for Slovak Music Economy Register itself. We can send all well-documented heritage collection items via Slovak Music Economy Register to Europeana and its more research-focused database.

Slovak Music Economy Register is technically stored in a MariaDB database extended for graph format with Wikibase. Wikibase, the software that powers Wikidata, the world's largest open graph database, has been successfully used in the past decade to create interoperable library services (van Veen 2019; Bianchini, Bargioni, and Pellizzari di San Girolamo 2021; Sardo and Bianchini 2022), museum services (Pfeiffer and Gayo 2021), and archives (references), and creating interoperability among the institutional silos of these three types of heritage collections (Alexiev et al. 2020; Rossenova, Duchesne, and Blümel 2022). EU institutions themselves also use it (Fischer, Thill, and Angjeli 2020).

- The author and the data stewards mentioned above oversees data curation activities. We are open to, and inviting co-creators with adequate knowledge about textile and fashion history expertise or data stewardship knowledge.
- Data stewards are responsible for quality assurance of data linking, data modeling, documentation, and technically correct updates.
- Researchers and users are encouraged to report errors and contribute to improvements.

A.6.9 Preservation and Archiving

- Historical datasets are archived following the practices of Zenodo for long-term storage.
- Data formats are chosen to ensure long-term accessibility and interoperability.
- Metadata standards (e.g., CIDOC-CRM, DCTERMS, DataCite, and general FAIR principles) are adopted not only for interoperability but also for long-term preservation.

A.6.10 Review and Continuous Improvement

- This data publication is reviewed periodically to align with evolving best practices and technological advancements.

A.7 Ethics statement

Data in Brief's Guide for Authors contains detailed information on the ethical guidelines with which all authors must comply. In addition, we ask you to complete the relevant statement(s) below. Please delete those which are not relevant for your data.

If your work involved human subjects,** please include a statement here confirming that the relevant informed consent was obtained from those subjects, if the research was carried out in accordance with the Declaration of Helsinki, and if it includes the **Ethical committee approval and the protocol

A.8 Acknowledgements

Please outline the contributions of each co-author, using the categories listed on the Elsevier website page.

Daniel Antal: Conceptualization, Methodology, Software, Writing, Funding acquisition. Richard Demčák: Data curation, Writing – review and editing. **Matej Grochal**: Data curation, Conceptualization, Writing – review and editing. **Marián Jankovič**: Writing – review and editing. **Anna Márta Mester**: Data curation, Investigation, Writing – review and editing. **Anna Žilková**: Data curation, Conceptualization, Methodology.

This document has been prepared by OpenMusE project partners as an account of work carried out within the framework of the EC-GA contract no 101095295 (Open Music Europe 2023).

The Open Music Europe project has received funding from the European Union's Horizon Europe, research and innovation programme, under Grant Agreement No. 101095295*. 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.*

Any dissemination of results must indicate that it reflects only the author's view and that the Commission Agency is not responsible for any use that may be made of the information it contains.

This is an early draft for internal discussion. It is not ready to be used or submitted for revision.

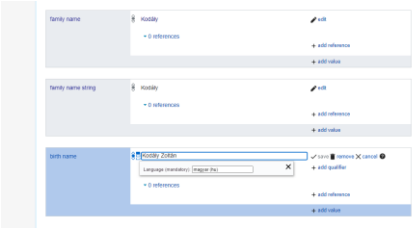
Appendix B — Appendix: How to use our registers

B.1 How to access our registers?

B.1.1 Human use

Please use for learning purposes our demo.

- Tutorial: SKCMDb import, 2024.05.21.
- Tutorial: Manual editing of persons' data in the shared register.



Please review our tutorial slides.

B.1.1.1 Examples: Persons, Corporate Bodies (Institutions & Companies) and Groups (Formal & Informal)

- composers in Slovakia: people who have composed music on the territory of the modern Slovak Republic, or composers of Slovakian origin. For example, Iris Szeghy has a data profile, which also connects here to the name authorities of ISNI (used on streaming platforms and in the future in collective management), VIAF (used by library catalogue systems), and of course, to her entry on the Slovak Music Center database. This way we want to cretea an interoperable register that allows the identification of composers and avoid ambiguity in royalty payments or attributions. Her body of work is also connected to his name.
- musical groups in Slovakia: musical groups on the territory of the modern Slovak Republic, or of performers of Slovakian origin. This is very important for rights management, because most uses of music is the use of sound recordings, which are manifestations of the compositions, and they are often identified by the performer, and not composer. To improve findability, we place the musical groups into subcollections, for example, choirs in Slovakia.
- music publishers in Slovakia: music pubilshing organisations (for example Hudobný Fond) and music publishing companies (for example A-Tempo Verlag s.r.o. Slovakia) that are registered with SOZA to collect royalties from Slovakia. The creation of the register of publishers is essential to monitor the market and create reliable statistics on the music industry.

B.1.2 API Access

API access will be provided, along with detailed instructions for use, via the Open Music Observatory curated by REPREX.

B.2 Who can access our registers?

We are creating shared registers to support parallel registration processes for different business, research, or governance uses. We utilise the novel data (sharing) space concept to harmonise data among registers and make some public registers available.

The answer to who can use these registers is “it depends on the use.” The natural persons register heavily relies on the collective management society’s membership and representation register in a given country (in our example, Slovakia.) This register can only be used for purposes that directly benefit the membership of such organisations. However, we encourage the membership to give us explicit consent to use their personal data for other purposes.

The use of our corporate body and primarily legal person (enterprise) registers is generally not restricted to any business, academic or policy research use, and they can support various applications like ethical recommender engines for streaming and background music or typically academic research and education.

The register of musical works is the property of SOZA or CISAC, and again, it cannot be used for any purposes outside of collective rights management for authors. The use of such data requires a contract with SOZA or CISAC.

The register of sound recordings and print music can be used for general purposes. We will demonstrate their use with improved public library services. The terms and conditions will be developed together with the first public use.