

Pilot Program for Novel Music Industry Statistical Indicators in the Slovak Republic

Business-to-government data sharing, novel re-use of public sector information
for the creation of missing marco-, industry-, and institutional KPIs for the
Slovak cultural and creative industry strategy implementation (working
document version 0.1)

Daniel Antal

2023-10-02

Table of contents

Executive Summary	2
Our approach to the problem	4
Data coordination	7
Slovak music industry enterprise register	8
Music professional population register	10
Ex ante survey harmonisation	12
Statistical production	13
Voluntary application of procedures	14
Application of open algorithms and open-source statistical code	15
Data stability: highly-standardised administrative data sources	15
Application of the open collaboration method	15
Retrospective harmonisation of survey data	16
Creation of novel indicators for the Slovak music industry and potentially other CCIs	16
Conclusion	17
References	17

Executive Summary

Access to high-quality official statistics is a public good that also increases the quality of decision-making on the enterprise and public policy levels. The problem of the cultural and creative sector institutions and enterprises is that they do not have access to the “music industry” or “film industry” statistics like banking or car manufacturing businesses do. They are less likely to engage effectively in the democratic tax policy-making process, for example. Cultural policy designers are also disadvantaged compared to their colleagues in tourism, agriculture, financial services or manufacturing.

Slovakia has already made a significant and exemplary investment in creating a satellite account system for the cultural and creative industries, which allows architecture and advertising to have high-quality data. However, the other creative industries can still only rely on more aggregated statistical indicators useful for macro-level public policy design but have minimal use for institutional or business policies. Our suggested cooperation aligns with the best European and UN statistical practices and recommendations and the data and digitisation strategies of the European Union and the Slovak government. They would enable the Ministry of Culture and the Institute for Cultural Policy to rely on evidence when adopting and implementing their public policies; it would also emancipate music businesses to rely better on the public good of high-quality indicators.

Open Music Europe offers a ‘data-to-policy’ pipeline, which extends the music data pipeline to evidence-based business and policy administration (Open Music Europe 2023a). A data pipeline is a method in which raw data is ingested from various data sources and then ported to data store, for further analysis, in this case, to an open, shared, collaborative music observatory. We extend this pipeline using reproducible research techniques, a novel application of the Open Policy Analysis Guidelines, and good statistical practices to support evidence-based policy analysis, scientific music research and sound business strategy building. In this last leg of the pipeline, we emphasise usability for our project’s target audiences and good documentation practices. We want to ensure that our data is high quality and well understood to support robust and correct business, scientific or policy conclusions.

Based on the *Memorandum o porozumení o využití výsledkov analýz otvorených politík v kontexte slovenského kultúrneho a kreatívneho priemyslu a sektorových verejných politík v spolupráci s konzorciom pre výskum a inovácie s názvom OpenMuse*. [Memorandum of Understanding on utilizing the Open Policy Analysis results of the OpenMuse Research and Innovation Consortium in the context of Slovak cultural and creative industries and sectors’ public policies, (Ministerstvo kultúry SR and Open Music Europe 2023)] we held our first stakeholder consultation in Bratislava on **21 September 2023**, where we invited representatives of the [Ministerstvo kultúry SR / Ministry of Culture of the Slovak Republic](#) and its Institute of Cultural Policy / [Inštitút kultúrnej politiky](#), the [Štatistický úrad SR / Statistical Office of the Slovak Republic](#) (henceforth: [SOSR](#)), and [Infostat](#), Institute of Informatics and Statistics, and representative Slovak stakeholders from the industry ([SOZA](#)) and from the music heritage sector and music libraries ([Hudobné centrum](#).)

 Tip

You can download the PowerPoint presentations and summaries of our stakeholder discussion with Dr Joost Poort (University of Amsterdam), Daniel Antal, CFA (Reprex), Dr. James Edwards (SINUS) and Magr. Tomas Miks (SOZA) [here](#).

With colleagues from IKP and the SOSR, we reviewed all the strategic indicators of the cultural, sustainable development and digitisation policy goals of Slovakia and the potential data sources of desired but not available policy indicators on macro-economic, industry, and institute/enterprise level. The creation of such KPIs is essential for the controlling of the policy execution and for the creation of transmission mechanisms to bring down the policy execution to the level of at least national organisations like SOZA or Hudobné centrum. The following document is detailed summary of the statistical aspects of the data coordination and collaboration presented in the workshop; we would like to discuss this approach on **10 October in Bratislava** with relevant statistical and music policy stakeholders.

We believe that our approach conforms not only with the relevant Slovak cultural, development, digitisation and statistical policies, but would constitute a best practice in terms of implementing and improving the creation of structural business indicators with the help of *privately-held data* (ESS 2017; HLEG 2020; European Commission 2020; ‘United Nations Guidelines on Statistical Business Registers. Final Draft Prior to Official Editing’ 2020).

The structure of this working document is as follows:

- [Our approach to the problem](#) in a nutshell conforming the case study and sandbox recommendations of the High-Level Expert Group on *European strategy on business-to-government data sharing for the public interest*, which is a supporting document of the *European Strategy for Data*, using good examples from various European countries (Norway is reviewed here.)
- [Data coordination](#) proposed by Open Music Europe among IKP, SOZA, Hudobné centrum, SOSR and the Open Music Europe Consortium.
- [Statistical production](#) of the novel indicators by voluntary application of official statistical guidelines and quality assurance, using advanced survey harmonisation, and application of the Eurostat indicator development methodologies.
- [Conclusions](#)
- [References](#) to the cited regulations, policy documents and academic literature.



ome Pilot Program for Novel Music Industry Statistical Indicators in the Slovak Republic

Open Music Europe – openmuse.eu – music.dataobservatory.eu – DOI: 10.5281/zenodo.8399254

Funded by the European Union

sinus: UNIVERSITY OF TURKU UNIVERSITY OF AMSTERDAM Institute for Information Law Sant'Anna School of Advanced Studies - Pisa UNIVERSITY OF ECONOMICS IN BRATISLAVA reprex SYNNO MUSIC INNOVATION HUB SOZA ALOADED meo MXF artisjus МУЗИКАТОР

Funded by the European Union under Grant 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 Commission's Citizens, Equality, Rights and Values Programme, or the official views of the Ministerstvo kultúry SR, Štatistický úrad SR, Infostat or IKP. Neither the European Union, the Slovak public bodies nor the granting authority can be held responsible for them.

Our approach to the problem

The music industry has no similar access to high-quality governmental statistics as a public good that SORS produces, such as banking or car manufacturing. We want to remedy the following problems with a proposed public-private partnership pilot project to improve the Slovak cultural and creative satellite accounts and generally the KPIs of the Slovak cultural strategy.

- State statistical survey programs under-sample the music sector, for example, in the Roč 1-01 and Roč 1-01 enterprise surveys or in LFS. The reason is the relatively small weight of the music sector in the Slovak economy, the fragmentation into small units, and the relatively small importance of the music professionals in the labour force. We want to remedy this situation by adding privately collected, ex ante harmonised questionnaire-based survey data to the existing Slovak surveys to reach a necessary joint dataset that allows statistical generalisation for music.
- State statistical survey programs use the NACE and ISCO classification systems to group enterprises and individuals, which are not suitable for the creative industries due to their high level of informality and the prevalence of informal learning instead of accredited education. We suggest the creation of the Slovak Music Industry Register, and a derived SBR satellite following UN/Eurostat methodology to improve

the identification of relevant data in the existing collection (music is covered in the collection but cannot be identified as “music”).

- The music industry, due to its high international standardisation and digitisation, has an unusually large number of well-maintained digital data resources that are not yet utilised in a way that credit card, supermarket, or energy distributor data is already used to create more granular and timely statistical products. Our Slovak Music Industry Register proposal would allow data coordination with well-standardised and quality-controlled music industry data.

In short, the statistical system does not have sufficient data about the music industry or the film industry; furthermore, it uses categorisation nomenclatures in the form of NACE and ISIC that are not suitable to filter out the existing data for most cultural and creative industries (CCIs) except architecture and advertising.

Several European countries are experimenting with new regulations that make data provision mandatory for certain enterprises like supermarkets, mobile telephone operators or credit card companies to provide privately-held, commercial data for official statistical purposes. The European System of Statistics (ESS), where SOSR is a member, had been advocating for the use of privately-held data in the last ten years.

Statistical surveys and administrative registers are and will remain important sources for the production of official statistics. However, the traditional data gathering methods could be in the future enhanced and enriched by big data analytics. To achieve this, it is essential that data held by private actors can be used by statistical authorities as raw material for innovative value-added services and statistical products, which will also boost the economy by creating new jobs and encouraging investment in data-driven sectors. Increased efficiency and faster delivery by statistical offices of innovative digital products and services will be one of the cornerstones for evidence-based decisions and contribute to lessening the burden on statistical respondents and notably on businesses. Improved access to data would enable statistical offices to provide more granular and timely statistics that would be more useful to enterprises and the citizens alike. It would also mean providing valuable feedback information to the data holders or delivering more tailored statistical services to companies, which might help them in return to better develop their business model. (ESS 2017, 3–4)

In preparation of several new data regulations within the European data strategy, the High-Level Expert Group on Business-to-Government Data Sharing presented an important and practical report under the title *Towards a European strategy on business-to-government data sharing for the public interest*. They use the B2G, or business-to-government abbreviations to new statistical procedures when official statisticians start to use private data; this is the opposite of the traditional “open data” regime, which allows the re-use of public sector information for the private sector.

The *High-Level Expert Group* advises that

B2G data-sharing collaborations should be organised:

- in testing environments (‘sandboxes’) for pilot testing (‘pilots’) to help assess the potential value of data for new situations in which a product or service could potentially be used (‘use cases’),
- via public-private partnerships. (HLEG 2020, 8.)

Member States should have in place structures to support B2G data sharing. These structures could be a body (or bodies) tasked with assisting public-sector organisations and private companies or civil-society organisations in entering into new data-sharing partnerships and facilitating the sharing of good practice. Over time, such structures could become trusted third parties between the public and private sectors, by bringing the relevant players together. (HLEG 2020, 37)

All B2G data sharing can ultimately be understood as a form of public-private partnership (PPP). However, from a practical point of view, sandboxes and specific PPPs can be put in place for specific public-interest purposes or to target specific challenges related to B2G data sharing. (HLEG 2020, p 41)

Our recommendation is to test in the sandbox of the **Open Music Europe** research and innovation action and project supported by the European Union, based on the *Memorandum of Understanding on utilizing the Open Policy Analysis results of the OpenMuse Research and Innovation Consortium in the context of Slovak cultural and creative industries and sectors’ public policies* (Ministerstvo kultúry SR and Open Music Europe 2023) to develop a use case in 2023-2025 that can inform interested partners by 2025 about potential public-private partnerships.

Sæbø and Dimakos review the statistical production methods for complication with privately held data, which provide insights for the data management of **Open Music Europe**, too. They also present the current state of affairs with the example of the Household Budget Survey (Sæbø and Dimakos 2023) which has been replaced by sampling privately-held data of credit card companies and supermarkets after decades of household surveying.

Until 2012, Norwegian statisticians measured household spending with a sample survey of Norwegians reporting their purchases; they asked a large sample of people drawn from the Norwegian population register (established in 1964) to fill in questionnaires. The data quality was particularly poor with food and beverages because such purchases, as opposed to infrequent buying of new television sets or festival tickets, happen daily; the reporting burden or the burden on memory is taxing when people retrospectively must fill out a questionnaire about buying bread, butter and orange juice.

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, p1.). 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 necessarily filters through the individual's cognitive biases and memory lapses. **Statistics Norway** realised that instead of asking 7000 households about what they were buying in the supermarkets, 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. CEEMID, the predecessor of **Open Music Europe**, relied on similar techniques that experimenting government statisticians used: we kept asking anonymously and randomly music creators about their received royalties, but we also compared the data with the actual anonymised payouts of Artisjus in Hungary and SOZA in Slovakia.

The lessons of Norway are interesting because the new statistical law (in force since 2021) allows such data collection after a cost/benefit analysis and risk reduction carried out by **Statistics Norway**. Norway, like all EFTA countries and Eurostat participates in ESS and applies the same statistical EU/EEA regulations that Slovakia, Bulgaria or Hungary applies. While in **Open Music Europe**, we are not planning and not even advocating state-mandated data collection; we find these criteria useful for voluntary data sharing with the government based on individual agreements, which we endorse in the music sector.

Data coordination

“A register aims to be a complete list of the objects in a specific group of objects or population.” (Anders and Britt 2007) We are planning music industry registers where the objects are *music works* and *sound recordings* (in statistical terms, music products), and the populations are *music authors*, *music performers*, *groupings of performers* (as the majority of the musicians perform, record, release in groups, ensembles, orchestras), *record labels* (which may be formal and informal businesses) and *music publishers* (enterprises.) From a statistical point of view, our planned music industry registers can be seen as “administrative registers” because they were not initially created for a statistical purpose by a statistical authority.

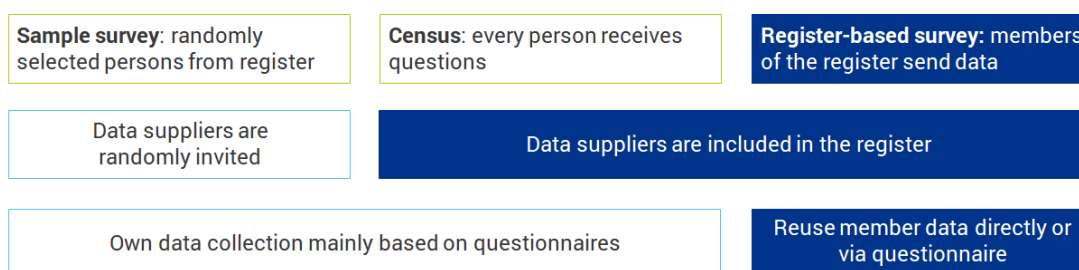
A *statistical register* is a continuously or regularly updated set of objects for a given population. It contains information on the identification and accessibility of population units as well as other attributes which support the surveying process of the population. It serves as a constantly updated 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

Guidelines on Statistical Business Registers. Final Draft Prior to Official Editing' 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 2021).

The statistical business register is an essential tool for creating survey frames or sample frames, in other words, to organise statistical data collection. In non-technical terms, this register is necessary to decide who should get a data request.

- For a **sample survey**, the register is used to draw a lottery of those members of the population who will be invited to provide data.
- In a **census*-type survey*, all registered members of the population, for example, all music labels, will receive an invitation to an interview or form.
- In the case of **a register-based survey**, all members of the register, for example, all collective management societies in the territory, will be requested to send data directly from their databases.



Slovak music industry enterprise register

The music industry needs a special register because the main nomenclature for categorising businesses, NACE, does not have a music industry entry; furthermore, only very few music enterprises participate in the state survey program because of their small size.

Most, but not all, music businesses are categorised as J58 (together with enterprises working for the film, television or radio industries) or R90 (together with dance, theatrical, circus and other performers, actors, and even film producers.) Any statistics calculated using the J58 or R90 category, for example, the total or average net turnover of such enterprises or their share in GVA and GDP, are only helpful for the macro-level cultural and creative industry policy design. The management of a music enterprise cannot compare their revenue, employee compensation or corporate profit and loss to such statistics because their changes in time do not only represent the business results of other music enterprises but unrelated film or theatrical revenues, too. For a public policy planner, a decrease in J58 is an early indicator of possible GVA decline and a reduction in tax receipts but gives no guidance if the problem is in the film, television or music industry. For the management of music enterprises, this indicator is far less helpful. When the total or average turnover decreases

or increases in J58, such a structural business statistical indicator change may be consistent with *both* decreasing music and growing film industry revenues.

We see a striking difference when comparing the information available for a music enterprise and a bank. Bank or insurance statistics are usually “clean” because engaging in these economic activities requires complex regulatory approval. An increase in turnover for K65 certainly shows an increased sales revenue for insurance companies, and a decrease in K64 average or total values only sets a different industry benchmark for non-insurance financial companies like banks and leasing firms. A bank executive or financial planner can take it for granted that the SORS will provide its team with relevant KPIs about demand for financial intermediation services, which will not be mixed up with insurance sales.

Creating the Slovak Music Industry Register should resolve this problem by defining whose answers should be asked for statistical purposes or whose responses should be filtered out or selected into music industry statistics. The register should contain the company registration numbers of entities that are active in a given business year in the Slovak music industry:

- **Music publishers** that register new works or receive royalties for previously registered (represented) works.
- **Record labels** that release new sound recordings or receive royalties for previously released recordings.
- **Concert and festival promoters, booking and talent agents, for-profit and social enterprises of musical groups** (ensembles, choirs, orchestras, bands) that create live music performances.
- Enterprises within the broader music industry are transport suppliers for music groups, producers or merchants of “merchandise”, music schools, rehearsal studios, and other enterprises.

Creating such an enterprise registry does not require particular data protection protocols. The music market is a public space, and legal entities must publicly account for their main economic activities. A good starting point is the enterprises that are registered with SOZA.

Most music enterprises are not providing statistical information to the SORS in any given year because only 2-3 Slovak music enterprises qualify for the enterprise size threshold for the mandatory Roč 1-01 survey. SORS randomly invites smaller music enterprises to participate in the Roč 2-01 survey not as “music enterprises” but as “sound recording” or “artistic performance” companies, undivided from companies working for film, television, or theatre. The creation of a music industry register enables the filtering of “music-only” responses in the survey and the calculation of statistical measures like “total revenue for the music industry” or “average number of employees in the music industry.”

In the banking industry, the registry is unnecessary because all banks go through regulatory approval and end up with the K64 code, and only large legal persons can provide banking services. Architecture is the most regulated in the cultural and creative industries: architects must usually meet strict certification criteria and register with a special chamber. No wonder

architecture is one of two CCIs covered *on industry level* by the Slovak Cultural and Creative Industry Satellite Accounts.

Because of their size, all banks will participate in the Roč 1-01 survey, which is the basis of the public structural business indicators concerning the financial benefits. With the help of their corporate registration number, the official statisticians can “pull” their financial data from the official company registry or their tax declarations, which are both critical “administrative data sources” for the SCSR. We cannot do this in the music industry because only a few music enterprises (as members of the J58 or R90 class) get an invitation to participate in the Roč 2-01 sample survey. The Slovak Music Industry Register serves too purposes in this case: it helps filter out the music industry-relevant answers from J58, R90 or other classes. It allows voluntary collection by the industry of further filled-in Roč 2-01 survey forms to meet the sample size requirement for music-industry statistics.

We discuss this problem and the proposed solutions at great length in our working paper, *Economy of music in Europe: Novel data collection methods and indicators* (Antal, Kmety Barteková, and Remeňová 2023) [see also: open policy analysis repository (Open Music Europe 2023b)].

Music professional population register

In theory, creating a registry of persons who engage professionally in music activities is not more complex than registering legal persons. The larger population size, the special protection of data and the privacy of individuals pose challenges in creating and using the population register in practice.

What type of natural persons should be part of the Slovak Music Industry Register?

- **Authors** who register without a publisher and self-publish music works and music lyrics.
- **Creators of sound recordings**, i.e., self-releasing producers and performers of recorded fixations of music performance that are released to be bought, streamed, broadcasted or played in the public against payment. Private persons who are *not* represented by a legal person music label.
- **Concert and festival promoters, booking agents, solo performers** and members of musical groups (ensembles, choirs, orchestras, bands) that create live music performances without being employed in a music institution or enterprises as a legal person.
- **Music professionals** that work within the broader music industry as suppliers of transport for music groups, producers or merchants of “merchandise”, music schools and rehearsal studios and other enterprises, as sole proprietors or in other forms that do not constitute a form of employment in a legal person.

National collective management societies like SOZA, SLOVGRAM or OZIS have a comprehensive list of creators who register their intellectual property for future payments or receive payments because certain forms of such revenues, like radio broadcasting royalties, are only paid by these organisations. The creators they represent individually or as members form the economically more active part of the music industry population. Experience of these organisations shows that some groups of music creators, particularly young creators in music genres that are not radio-friendly because of their eccentric music or explicit lyrics, fail to register with these organisations because they do not expect broadcasting royalties. National collective management societies must try to attract such creators to their member base because the copyright management infrastructure offered by such organisations goes beyond the administration of the royalties that may not be available for such artists.

Creators tend to be the minority of the music workforce. Behind a solo artist or a band of four, we usually find at least five music professionals behind the stage who manage the sound, light, stage technology, the administration and sale of tickets and merchandise, transport and install the technology, or control the crowd at entry, exit and in emergency. In more prominent stages and larger ensembles, the support crew size expands to 200 people, while even a symphonic orchestra is made of “only” 80-100 performers. Although our earlier research shows that within the music ecosystem, individuals usually perform more than one activity (they promote concerts and perform music or act as both sound engineers and music creators), a very sizable part of the music industry population has no copyrights or neighbouring rights and remains invisible for the collective management societies.

The natural person population of the music industry appears in population censuses every ten years and randomly in sample surveys, like the Labour Force Survey (LFS) or the Adult Education Survey (AES), which is conducted in Slovakia and every EU member state in a harmonised manner. The problem of these sample surveys is similar to the enterprise surveys: it is based on a categorisation, ISCO, which does not allow effective filtering out of responses given by the people working or learning in the music industry. Because of the relatively small size of the music sector, it is possible (but far from certain) that there are not enough responses recorded in the SOSR surveys to create statistics about the music industry; unfortunately, even if the sample size is large enough, it is impossible to select those responses that relate to music.

In the case of natural persons, we suggest the creation of ex-ante harmonised surveys, which ensure that enough music professional persons give answers to questions that a few anyways respond as members of a random sample in LFS or AES. The coordination of these surveys is nevertheless more complicated because we cannot simply use a natural person identifier as a company registry number to file or join data. Yet, the creation of the population register is still a first step. Before we define who belongs personally to the music industry, it is not possible to administer a survey that gauges their opinions or characteristics.

In the past years, SOZA, in methodological cooperation with Artisjus in Hungary and HDS in Croatia, carried out sample surveys in the music professional population that achieved a high level of representativeness among creators, even though these surveys were not based on a music professional population register. At the same time, they did not know who filled out the CEEMID music professional surveys; two mechanisms secured

representativeness. Even though in the absence of a register, explicit stratification was not possible with the survey, SOZA (and Artisjus) could ascertain that the entire population receives a questionnaire. The form remained anonymous, but the whole population was invited to fill it out. Furthermore, as SOZA (and Artisjus) administers exclusively and comprehensively particular types of royalties, we could compare the true statistics (mean, median, standard deviation) of actual royalty payments to anonymously reported amounts. We collected answers until the anonymous sample's statistics converged to the known, true statistical values in these anchoring points.

This leads to the third part of the use of the register: access to administrative data, i.e., demographic and income information from comprehensive records of IT systems that were not designed primarily for statistical production but for the administration of music royalties (SOZA, SLOVGRAM, OZIS), or the documentation of music heritage (Hudobné centrum and the Slovak music libraries.)

Ex ante survey harmonisation

Retrospective survey harmonisation usually refers to social science surveys conducted with a questionnaire, when researchers expose randomly selected respondents to randomly assigned treatments—for example, ask the respondent about their subjective well-being before and after taking a pill or a placebo and combining it with blood pressure or weight measurement. These measurements may be human-transcribed to the survey questionnaire or recorded by a different tool when a new data integration problem occurs.

Our researchers have long been engaged in retrospective survey harmonisation, for example, in the case of Cultural Access and Participation surveys with the methodology created by the ESSNet-Culture working group of Eurostat and the participating EU national statistical authorities (Haan and Adolfsen 2008; Bina, Vladimir et al. 2012; Haan and Broek 2012). We have extensive experience administering CAP surveys in Slovakia and Hungary and retroactively harmonising them with CAP surveys carried out within various EU-harmonised survey programs, such as Eurobarometer, EU-SILC and AES. Retrospective survey harmonisation can join data from different surveys if they use a similar sampling method and questionnaire items. If ex ante harmonisation is possible before the fieldwork, a much higher quality of harmonisation is possible.

“To ensure that answers from respondents surveyed in different settings carry minimal methodological errors and biases and can be meaningfully compared, both data producers and secondary users combine surveys from different sources, that is, they harmonize survey data. Generally, they do so at different stages of the survey lifecycle. Data producers mostly employ harmonization ex-ante, when designing and implementing comparative studies (input harmonization) and when processing the survey data in preparation for their public release (ex-ante output harmonization). [...] Secondary users apply harmonization methods retrospectively to already released data files.” (Wysmulek, Tomescu-Dubrow, and Kwak 2022)

We would like to ex ante harmonise additional surveys to meet the subsample requirements for music professionals and music enterprises in Roč 1-01 and Roč 2-01; or, if this is not possible, to use “small area statistics” to work from smaller samples.

Statistical production

The methodological framework that **Statistics Norway** is preparing has three interlinked pillars:

- Data minimisation by sampling
- Improving confidentiality
- Minimisation of data storage.

After consultation with the data protection authority of Norway, **Statistics Norway** concluded that sampling should occur outside the statistical office, i.e., before the data leaves the original data controller, for example, a collective rights management organisation. The **Open Music Europe** consortium members hold a similar view with a slightly different data governance angle. In our case, we are dealing with strictly voluntary data provision for the statistical authorities. Collective management organisations can only maintain their duty to their members if they do it themselves and guarantee the reduced risks of data sharing. Of course, we must consider the differences of the underlying data; in the Norwegian example “nano-level” transnational data is considered; in our case, we are considering microdata coordination and harmonisation.

Statistics Norway also experienced new data quality and reproducibility issues that **Open Music Europe** faces. It had been unprecedented in the production of official statistics that crucial parts of the statistical production or quality control are “outsourced”, but in the case of privately-held data, this is necessary for data protection. (From the viewpoint of SOZA and other collective management societies, they are “insourcing” those statistical procedures that government statisticians are outsourcing due to data protection concerns.)

In this case, the problem arises that the privately held organisations incur significant data processing costs. In the case of **Open Music Europe**, these costs are recovered from a public European Union grant, where data programming, coordination, and management are paid for by SOZA, UTU and Reprex. There is a lively debate on the European level on how and to what extent national (official) statistical budgets should reimburse the costs of privately-held data providers. Filling out census or sample surveys is mandatory. Citizens and institutions usually pay with their time or their employees’ time when they fill out these questionnaires or answer interviews. There is a consensus that participation in statistical interviews, for example, should not be incentivised with payments. These arguments cannot be held in the case of a few select private institutions expected to in-source costly statistical production and quality control measures necessary for trustworthy data provision.

Open Music Europe aims to solve these problems with the introduction of several novelties:

- The voluntary applications of specific guidelines originally designed for government statisticians;
- The use of open-source statistical software that applies statistical methods and procedures in a way that satisfies the needs of official statisticians;
- The use highly standardised, stable datasets;
- The use of open data management practices in the open music observatory, which design and document procedural safeguards for data quality.

Voluntary application of procedures

Our data management plan and practices aim for high-level compatibility and interoperability with official statistics. We use the definitions from the appropriate ESS glossaries and regulations whenever possible and try to apply (with the necessary modifications).

- We follow the definitions from Article 2 of the *2019/1700 EU Regulation on establishing a common framework for European statistics relating to persons and households, based on data at individual level collected from samples*, the definitions of the *2016/679 EU General Data Protection Regulation*, and the Article 3 of *223/2009 EC Regulation on European statistics* (European Parliament and Council of the European Union 2019, 2009; European Parliament and the Council 2016; EUR-Lex 2016, 2015). While the latter regulation is no longer in force, its definitions became parts of the European statistical vocabulary. For administrative data we use the [Cross-linked glossary for ADMIN data](#) in the *ADMIN Knowledge repository* of the European System of Statistics.
- The *European Statistics Code of Practice* and the *Quality Assurance Framework of the European Statistical System* (European Statistical System Committee 2018; European Statistical System 2019) are aimed at national statistical authorities. We aim to apply their provisions wherever possible, with the necessary difference in interpretation for state-mandated and voluntary data collection.
- On the level of data coordination, we follow the recent statistical guidelines on using privately-held data to create official statistics (UNECE 2015; ‘United Nations Guidelines on Statistical Business Registers. Final Draft Prior to Official Editing’ 2020; Eurostat 2021) , and, in broader terms, the management and policy literature of business-to-government data sharing (ESS 2017; HLEG 2020; Vigorito 2022; Sussha et al. 2022).
- We thrive to build a best policy practice in Slovakia within the context of the *European data strategy* (European Commission 2020).

Application of open algorithms and open-source statistical code

Within the framework of rOpenGov, UTU and Reprex are developing extensions (software packages) to the R statistical environment and language. National statistical offices widely use R in statistical production. Even if they do not use it, R is a high-level, interpreted language that production statisticians can read, even if they are used to working with different computer languages. To provide the necessary quality control of the statistical production pipeline, the data providers (in Slovakia, initially SOZA) will only use open-source statistical code that government statisticians can vet. The researchers of UTU and Reprex will send all key elements of their software to peer review first on CRAN, then within the broader data science and academic statistical community. This measure will provide a necessary safeguard to the production quality and reproducibility regarding the production itself.

Data stability: highly-standardised administrative data sources

Another safeguard of the reproducibility concerning the data sources is the relative stability of collective management data. Music licensing is a very highly standardised global activity. Royalty accounting music metadata standards ensure that SOZA can cooperate and exchange information about distributing money to rightsholders in almost every country. Global organisations like CISAC and BIEM govern these standards, often implementing stable international standards or agreements. Data stability, another critical aspect of statistical reproducibility within a state statistical program, can be considered better founded than in the case of many public administrative records that SOSR or other national statistical offices have been using in the past decades.

Application of the open collaboration method

We practice the open collaboration method, rooted in available knowledge management (library science) and open-source software development. As long-time developers of open-source statistical software on the rOpenGov platform, we use good project management practices, code sharing and data sharing to facilitate work across business, scientific, and government entities and allow the participation of individual researchers, developers or citizen scientists, too.

- We have adopted the *Open Policy Analysis Guidelines* to make our work easy to follow, review and connect to (Hoces de la Guardia, Grant, and Miguel 2021). This means that like open source software developers would do, we make our policy analytic work, including this document and its supporting reader, references, charts, available for review, criticism and improvements in a standard file structure on the GitHub open repository.
- The software library extensions to the R statistical environment and language are available in development version, daily refreshed in the GitHub repositories of [rOpenGov](#)

managed by the University of Turku Data Science Group; and after necessary peer-review and tasting on the [The Comprehensive R Archive Network](#).

- All our developers and researchers must adhere to the [Contributor Covenant](#) stewarded by the Organization for Ethical Source.

Retrospective harmonisation of survey data

Creation of novel indicators for the Slovak music industry and potentially other CCIs

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

We reviewed three Slovak policy documents for our work which are relevant: the sectorial *Stratégia kultúry a kreatívneho priemyslu Slovenskej republiky 2030* (Ministerstvo kultúry Slovenskej republiky 2023)—henceforth: Slovak CCI Strategy, the *Slovakia’s Vision and Development Strategy 2030 - a long-term strategy for sustainable development* (Ministerstva investícií, regionálneho rozvoja a informatizácie Slovenskej republiky 2020)—henceforth: Slovak Strategic SDGs, which has significant cultural elements, and the *2030 Digital Transformation Strategy for Slovakia* (Ministerstva investícií, regionálneho rozvoja a informatizácie SR 2019a, 2019b)—henceforth: Slovak Digitisation Strategy, which is highly relevant for the innovation we are doing in Slovakia under the label “Listen Local” in WP2.3. These documents were made in different years, and they are consistent with the EU policy document of that given year; they show little cross-referencing however within the Slovak documents. This means that for example, the most dated Slovak Digitisation Strategy is having no reference to the cultural and creative industries at all, but also, the later Slovak CCI Strategy has very little concrete reference to the SDG policies and no reference to the Slovak Digitisation Strategy. The Slovak CCI Strategy is a well-designed policy paper, and it is easy to link to the relevant, horizontal SGD and digitisation policies.

Our literature and data source review is building on the earlier work of KEA European Affairs, who, in cooperation with the European Commission’s Structural Reform Support Service (SRSS), reviewed the policy documents preceding the final version of the Slovak CCI Strategy, and almost all data sources that are relevant for our work, too (KEA 2020).

The “triple transition” is highly relevant to our work because these horizontal policies guide public investment, have new rules on private investment spending, and foresee changes in economic and tax policies. It immediately connects with the 2nd strategic priority of the Slovak CCI Strategy, 2 Efficiently funded culture to systematically reduce the infrastructure and modernisation gap, increase the efficiency of the finance management and financing of culture and creative industries, and complement public funding sources with private sources. Furthermore, we also agree with the strategy’s prioritisation to put 3 Dignified culture, i.e.,

the proper remuneration for the workforce of the Slovak cultural sectors, on the top of the policy intervention agenda, because in CCSIs, the bigger part of the gross value added or gross product is the value created by the workforce (labour) and not fixed capital assets. The following KPIs of this strategic objective are very relevant to our work.

The indicators from the text below are added to the slides, too.

The Slovak strategic SDGs have a strong cultural component, too (Ministerstva investícií, regionálneho rozvoja a informatizácie Slovenskej republiky 2020, 58.), which partly overlap with the Slovak CCI Strategy.

Conclusion

References

- 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, 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>.
- Bína, Vladimír, Chantepie, Philippe, Deboin, Valérie, Kommel, Kutt, Kotynek, Josef, and Robin, Philippe. 2012. 'ESSnet-CULTURE, European Statistical System Network on Culture. Final Report.' Edited by Frank, Guy. https://ec.europa.eu/assets/eac/culture/library/reports/ess-net-report_en.pdf.
- 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>.
- EUR-Lex. 2015. *Consolidated Text: Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European Statistics and Repealing Regulation (EC, Euratom) No 1101/2008 of the European Parliament and of the Council on the Transmission of Data Subject to Statistical Confidentiality to the Statistical Office of the European Communities, Council Regulation (EC) No 322/97 on Community Statistics, and Council Decision 89/382/EEC, Euratom Establishing a Committee on the Statistical Programmes of the European Communities (Text with Relevance for the EEA and for Switzerland)*. <http://data.europa.eu/eli/reg/2009/223/2015-06-08/eng>.
- . 2016. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA Relevance)*. <http://data.europa.eu/eli/reg/2016/679/oj>.
- European Commission. 2020. '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>.

- European Parliament and Council of the European Union. 2009. *Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European Statistics and Repealing Regulation (EC, Euratom) No 1101/2008 of the European Parliament and of the Council on the Transmission of Data Subject to Statistical Confidentiality to the Statistical Office of the European Communities, Council Regulation (EC) No 322/97 on Community Statistics, and Council Decision 89/382/EEC, Euratom Establishing a Committee on the Statistical Programmes of the European Communities (Text with Relevance for the EEA and for Switzerland)*. <http://data.europa.eu/eli/reg/2009/223/oj/eng>.
- . 2019. ‘Regulation (EU) 2019/1700 of the European Parliament and of the Council of 10 October 2019 Establishing a Common Framework for European Statistics Relating to Persons and Households, Based on Data at Individual Level Collected from Samples, Amending Regulations (EC) No 808/2004, (EC) No 452/2008 and (EC) No 1338/2008 of the European Parliament and of the Council, and Repealing Regulation (EC) No 1177/2003 of the European Parliament and of the Council and Council Regulation (EC) No 577/98 (Text with EEA Relevance)’. <http://data.europa.eu/eli/reg/2019/1700/oj/eng>.
- European Parliament and the Council. 2016. *Consolidated Text: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA Relevance)*. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02016R0679-20160504&qid=1688020749152>.
- European Statistical System. 2019. ‘Quality Assurance Framework of the European Statistical System. 2019 Edition (Version 2.0)’. <https://ec.europa.eu/eurostat/documents/64157/4392716/ESS-QAF-V1-2final.pdf/bbf5970c-1adf-46c8-afc3-58ce177a0646>.
- European Statistical System Committee. 2018. *European Statistics Code of Practice — Revised Edition 2017. English Version*. Luxembourg: Publications Office of the European Union. <https://ec.europa.eu/eurostat/web/products-catalogues/-/ks-02-18-142>.
- Eurostat. 2014. *Towards a Harmonised Methodology for Statistical Indicators — Part 1: Indicator Typologies and Terminologies*. 2014th ed. Vol. 1. Towards a Harmonised Methodology for Statistical Indicators 1. Luxembourg: Publications Office of the European Union. <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-14-011>.
- . 2017. *Towards a Harmonised Methodology for Statistical Indicators — Part 3: Relevance for Policy Making*. 2017th ed. Vol. 3. Towards a Harmonised Methodology for Statistical Indicators 1. Luxembourg: Publications Office of the European Union. <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-17-007>.
- Eurostat. 2021. *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>.
- Haan, Jos de, and Anna Adolfsen. 2008. *De Virtuele Cultuurbezoeker - Publieke Belangstelling Voor Cultuurwebsites*. SCP-Publicatie 2008/9. Den Haag, the Netherlands: Sociaal en Cultureel Planbureau. <https://archieff18.archiefweb.eu/archives/archiefweb/20200311100055/https://www.scp.nl/dsresource?objectid=>

- 5c6903f1-e6ae-4e62-9a9b-6d5e1529756a&type=org.
- Haan, Jos de, and Andries van den Broek. 2012. ‘Nowadays Cultural Participation - an Update of What to Look for and Where to Look for It’. In *ESSnet-CULTURE, European Statistical System Network on Culture. Final Report.*, 397–417. Luxembourg. https://ec.europa.eu/assets/eac/culture/library/reports/ess-net-report_en.pdf.
- HLEG. 2020. *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.* Luxembourg: Publications Office of the European Union. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=64954.
- Hoces de la Guardia, Fernando, Sean Grant, and Edward Miguel. 2021. ‘A Framework for Open Policy Analysis’. *Science & Public Policy* 48 (2): 154–63.
- KEA. 2020. ‘Cultural Industries in Slovakia. Report on the AS-IS Situation. Analysis and Recommendations to Boost Growth in Cultural Industries in Slovakia’. Brussels: KEA European Affairs. https://www.culture.gov.sk/wp-content/uploads/2020/11/AS-IS_sprava.pdf.
- Kotzeva, Mariana, Anton Steurer, Nicola Massarelli, and Mariana Popova, eds. 2017. *Towards a Harmonised Methodology for Statistical Indicators — Part 2: Communicating Through Indicators.* 2017th ed. Vol. 2. Towards a Harmonised Methodology for Statistical Indicators 1. Luxembourg: Publications Office of the European Union. <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-17-001>.
- Ministerstva investícií, regionálneho rozvoja a informatizácie Slovenskej republiky. 2020. ‘Vízia a stratégia rozvoja Slovenska do roku 2030 - dlhodobá stratégia udržateľného rozvoja [Slovakia’s Vision and Development Strategy 2030 - a Long-term Strategy for Sustainable Development]’. <https://mirri.gov.sk/wp-content/uploads/2021/01/Slovensko-2030.pdf>.
- Ministerstva investícií, regionálneho rozvoja a informatizácie SR. 2019a. ‘2030 Digital Transformation Strategy for Slovakia’. Ministerstva investícií, regionálneho rozvoja a informatizácie SR. <https://mirri.gov.sk/wp-content/uploads/2019/10/SDT-English-Version-FINAL.pdf>.
- . 2019b. ‘Stratégia digitálnej transformácie Slovenska 2030. [20230 Digital Transformation Strategy for Slovakia]’. Ministerstva investícií, regionálneho rozvoja a informatizácie SR. <https://mirri.gov.sk/wp-content/uploads/2019/06/Strategia-digitalnej-transformacie-Slovenska-2030.pdf>.
- Ministerstvo kultúry Slovenskej republiky. 2023. ‘Stratégia kultúry a kreatívneho priemyslu Slovenskej republiky 2030. [Strategy of the cultural and creative industries of the Slovak Republic 2030.]’. https://strategiakultury.sk/wp-content/uploads/2023/06/Strategia_KKP_SR_2030_po_shvaleni_.pdf.
- Ministerstvo kultúry SR, and Open Music Europe. 2023. ‘Memorandum o porozumení o využití výsledkov analýz otvorených politík v kontexte slovenského kultúrneho a kreatívneho priemyslu a sektorových verejných politík v spolupráci s konzorciom pre výskum a inovácie s názvom OpenMuse. [Memorandum of Understanding on utilizing the Open Policy Analysis results of the OpenMuse Research and Innovation Consortium in the context of Slovak cultural and creative industries and sectors’ public policies]’. <https://www.crz.gov.sk/zmluva/7645338/>.
- Open Music Europe. 2023a. ‘Open Music Europe (OpenMusE) – An Open, Scalable, Data-to-Policy Pipeline for European Music Ecosystems’. <https://doi.org/10.3030/>

- 101095295.
- . 2023b. ‘Repository: Report on the European Music Economy. GitHub’. Open Policy Analysis Repository with Standardized File Structure. 26 March 2023. <https://github.com/dataobservatory-eu/report-european-music-economy>.
- Sæbø, Hans Viggo, and Xenia 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>.
- Susha, Iryna, Jakob Schiele, Koen Frenken, Marijn Janssen, Csaba Csáki, Ida Lindgren, Ulf Melin, et al. 2022. ‘Business-to-Government Data Sharing for Public Interests in the European Union: Results of a Public Consultation’. In *Electronic Government*, 13391:515–29. Lecture Notes in Computer Science. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-15086-9_33.
- 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 Guidelines on Statistical Business Registers. Final Draft Prior to Official Editing’. 2020. United Nations. https://unstats.un.org/unsd/business-stat/SBR/Documents/UN_Guidelines_on_SBR.pdf.
- 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* 9 (3): 237–58. <https://doi.org/10.1163/22134514-bja10030>.
- Wysmulek, Ilona, Irina Tomescu-Dubrow, and Joonghyun Kwak. 2022. ‘Ex-Post Harmonization of Cross-National Survey Data: Advances in Methodological and Substantive Inquiries’. *Quality & Quantity* 56 (3): 1701–8. <https://doi.org/10.1007/s11135-021-01187-7>.