

Final report

Good practices in accessing, using and contributing to the management of administrative data

Reference period	2015-2017 (13/11/2015 to 12/5/2017)
Date of delivery	24/09/2018
Contractor	Agilis SA

Document service data

Type of Document:	Project Deliverable		
Contract Full Title:	Legal and institutional framework for the use of administrative data (ESS.VIP.BUS.ADMIN WP1 Finalisation of the draft reports)		
Contract Number:	07112.2015.004 – 2018.0141		
Version:	3.0	Status:	Final
Created by:	Anais Santourian, Stella Kitromilidou, Gregory Farmakis	Reviewed by:	Anais Santourian
Distribution:	Eurostat Unit F2, Agilis S.A.		

Document History

Version	Date	Comments
1.0	06/07/2018	Revised version based on comments received from countries
2.0	13/09/2018	Revised version based on the comments from the 2 nd round of consultation with countries
3.0	24/09/2018	Final version

The information and views set out in the report are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

Table of contents

1. Introduction	6
1.1. Data sources.....	7
2. Good Practices and Recommendations	8
2.1. Legal Frameworks	8
2.1.1. Statistical Laws	8
2.1.2. Laws governing data sources.....	10
2.2. Institutional and Organisational Aspects	10
2.2.1. Internal coordination practices	10
2.2.2. Cooperation with other stakeholders.....	12
2.2.3. Communication with the public.....	13
2.3. Cooperation with Data Owners and Wider Institutional Environment	14
2.3.1. Establishing the cooperation	14
2.3.2. National Data Sharing Strategies	16
2.3.3. Interoperable national infrastructures.....	17
2.3.4. Synergies with e-government applications and digital services.....	17
2.3.5. Other Cooperation Aspects	17
2.4. Quality and Technical Aspects	18
2.4.1. Methodological Work	18
2.4.2. Exploitation of technological standards	19
2.4.3. Extended Applications	20
3. Examples of data access cases	21
3.1. Data sources.....	21
3.2. Success factors and obstacles	23
4. Concluding Remarks.....	28

Abbreviations

EFTA	European Free Trade Association
ESS	European Statistical System
ESSnet	Collaborative ESS networks
ESS.VIP ADMIN	ESS Vision 2020 Implementation Project for Administrative Data
EU-SILC	European Union Survey on Income and Living Conditions
EU-LFS	European Labour Force Survey
MS	Member States
NSI	National Statistical Institute
SBS	Structural Business Statistics
STS	Short-term Business Statistics
WP	Work Package

1. Introduction

The world of statistical data is evolving rapidly. Constantly evolving data needs, the emergence and proliferation of non-traditional data sources, and an enabling technological environment introduce a new era. With pressures mounting due to surging demands for more data and unfavourable budgetary conditions, administrative data figure prominently among promising new initiatives. While the issue of administrative data as such is not new, it has had different histories and has enjoyed more traction in some countries than others. Moreover, administrative data have several dimensions ranging from the legal and cultural to the practical and technological, and concrete progress depends on the intersection of them all.

Eurostat supports the quest of NSIs to make serious inroads in this area, including learning from each other under the framework of the Work Package 1 (WP1) of the ESS.VIP ADMIN¹ (Administrative data sources) project. ESS.VIP ADMIN is one of the projects that contributes to implementing the European Statistical System Vision 2020².

The work undertaken under the framework of WP1 is reflected in the present report on best practices regarding access to data and relation of the NSIs with the administrative data owners which complements the report on the legal and institutional environment in MSs.

This report is one of the outcomes of the task “Legal and institutional framework for the use of administrative data (ESS.VIP ADMIN WP1, Access to and development of administrative data sources)”³ and aims to contribute to this effort. Its purpose is to highlight good practices identified in accessing administrative data. It does so by adhering to the same four-dimensional classification developed and utilised in the report “Analysis of the legal and institutional environment in the EU Member States and EFTA countries”. That is, issues are grouped under four thematic dimensions: (1) legal frameworks; (2) institutional and organisational aspects; (3) cooperation with data owners and wider institutional environment, and; (4) quality and technical aspects.

In addition to these dimensions, national differences in administrative data exist along specific groupings, such as data for individuals or businesses or more specifically across specific administrative sources, e.g. taxation, health, education, business registers etc. Therefore, this report also includes relevant findings concerning examples of exemplary and problematic data access cases at that level across individual countries.

It must be emphasised from the outset that the countries mentioned are used as examples of some specific good practice. This does not mean that the countries named are the only examples of that particular good practice nor that the examples used are necessarily the best. All examples are for illustration purposes only, and there may well be NSIs in other countries that have not been covered. A key reason for this is that emphasis has been on identifying examples in the framework of the information gathering exercise undertaken for this project rather than on standardizing data sources and statistical domains across countries or compiling an exhaustive or ranked list.

Although it is hoped that this report will serve as a useful stock-taking exercise and will contribute to the ongoing discourse and cross-fertilisation among NSIs, it should be read with these qualifications and caveats in mind.

¹ https://ec.europa.eu/eurostat/cros/content/ess-vision-2020-admin-administrative-data-sources_en

² <http://ec.europa.eu/eurostat/web/ess/about-us/ess-vision-2020>

³ “Legal and institutional framework for the use of administrative data (ESS.VIP ADMIN WP1, Access to and development of administrative data sources)” (Eurostat, Contract No. 07112.2015.004-2015.678).

1.1. Data sources

In the framework of the data collection exercise undertaken for the needs of the project (see also the report “Analysis of the legal and institutional environment in the EU Member States and EFTA countries”), an extensive review of the information available in public documentation, Admin Data ESSnet deliverables⁴ and the national peer review reports⁵ were examined.

A data collection exercise was launched during the period August – October 2016. The NSIs were invited to fill-in two questionnaires, namely the “Strategy level questionnaire” and “Operational level questionnaire”, with the view to collect additional information by the NSIs on the obstacles faced in accessing administrative data as well as best practices for overcoming these obstacles.

Overall, 30 replies were received to the Strategy level questionnaire and 27 to the Operational level questionnaire⁶.

Finally, a Workshop on access to administrative data sources was held in Brussels, on 13-14 September 2016⁷. The aim of the Workshop was to facilitate information exchange and sharing of experiences among NSIs’ representatives.

In all cases where examples are mentioned, the sources of information were (a) the national peer review reports, (b) the replies of the NSIs to the questionnaires and (c) the report of the Workshop on access to administrative data sources⁸.

⁴ https://ec.europa.eu/eurostat/cros/content/use-administrative-and-accounts-data-business-statistics_en

⁵ <http://ec.europa.eu/eurostat/web/quality/peer-reviews>

⁶ Strategy level questionnaire: Replies received by BE, BG, CZ, DK, EE, IE, EL, ES, FR, HR, IT, CY, LV, LT, LU, HU, MT, NL, AT, PL, PT, RO, SI, SK, FI, SE, IS, NO, CH. Operational level questionnaire: Replies received by: BE, BG, CZ, EE, IE, EL, ES, HR, IT, CY, LV, LT, LU, HU, MT, NL, AT, PL, PT, RO, SI, SK, FI, SE, UK, IS, NO, CH.

⁷ https://ec.europa.eu/eurostat/cros/content/2016-workshop-access-administrative-sources_en

⁸ https://ec.europa.eu/eurostat/cros/system/files/adda-workshop_technical_report_0.pdf

2. Good Practices and Recommendations

The focus of this chapter is on good practices in the acquisition and use of administrative data that were gathered during this project. Numerous good practices are identified and discussed, together with further potential improvements that could be conducive to the continuous incorporation of administrative data in official statistics.

The structure of the chapter follows the four-dimensional classification outlined in the introduction. Under legal framework, strong legal provisions in the national statistical laws for the NSI's right to access and use administrative data sources are discussed. This is followed by examples of good practices related to institutional and organisational aspects both within NSIs and between NSIs and other stakeholders. Then, examples of good practices in the cooperation with data owners are offered. Lastly, the discussion extends to interesting practices in conjunction with data quality and technical aspects.

2.1. Legal Frameworks

It is widely accepted that strong legal provisions that enable NSIs to acquire administrative data and use them for statistical purposes lay the foundation for the terms of conduct with data owners and frame the building of relationships in this area. Much progress has been made in recent years in terms of enacting new laws or strengthening the provisions of existing ones. The remainder of this section highlights examples of good practices at this point in time.

It should be noted that many of the following factors are not in the direct control of the NSI. Nevertheless, the NSI may influence them by actively participating in the public debate or formal deliberations preceding new legislation or even by communicating priorities to higher levels of the government and politicians.

2.1.1. Statistical Laws

The explicit provision in the relevant statistical law for the NSI's right (or even mandate) to access and use administrative data sources is a good practice to alleviate potential objections by data owners.

This was the case e.g. in Poland, where the statistical law has been revised to explicitly include provisions for access to administrative data. Moreover, as stipulated in Article 13.5 of the Law on Official Statistics: "while selecting the data sources for the purpose of official statistics, administrative data shall be considered in the first place". Similarly, in the UK the statistical law is under revision to give the NSI stronger data access rights. During the two-year public consultation of a proposed new statistical law that would give it stronger rights to access administrative data, the ONS was extensively involved in dialogue with politicians, data privacy lobbies, business associations, citizen groups and others. In 2015, the peer reviewers mentioned that in the context of deliberation for a new law on Danish Statistics, an inter-ministerial committee would be established to prepare the proposal for a new law and in that case Statistics Denmark would take active part in the work of this committee. Also, according to the peer review report (2015), Statistics Austria would communicate to the relevant authorities that the peer reviewers recommended an amendment of the statistical law, and would propose to launch a dialogue striving for an improvement of the legislation. Similarly, the peer reviewers stated in 2015 that Statistics Belgium would launch a dialogue striving to bring legislation at all relevant levels in line with the revised 223/2009 regulation, in particular with article 17a.

Beyond provisions for access rights, a stronger legal mandate is an efficient way to overcome objections or reluctance from data owners, or even internal cultural obstacles. Such a stronger

mandate might enforce the usage of administrative data by actually forbidding the carrying out of a survey if the desired data are already available in an administrative source.

Thus, in Germany, the new version (section 5a) of the Federal Statistics Act, integrated in the version promulgated in October 2016, is intended to encourage both the legislators and the statistical offices to envisage more often the use of administrative data for the production of federal statistics (peer review report, 2015). While administrative data are frequently used already now, the new regulation will make such use a general principle. The aim of the new regulation is that institutions owning administrative data will for the first time be obliged to formally cooperate with the Federal Statistical Office to enable the latter to obtain metadata of those administrative data. The new regulation will also include a provision stipulating that institutions owning administrative data are obliged to transmit, upon request, formally anonymised microdata from their stock of administrative data to the Federal Statistical Office so that the data can be examined for suitability.

The Austrian National Statistical System is obliged whenever possible to use administrative data sources in preference to direct surveys. Article 6 (3) of the Austrian Statistics Law states that surveys should only be conducted when data are not available from administrative databases, registers or other databases. Similarly, according to the provisions laid down in Section 4 of the Statistics Finland Act, the NSI has a strong mandate to use administrative data in preference to direct surveys for the compilation of statistics. The Statistics Lithuania's policy prescribes that data from respondents should be collected only if administrative or other statistical data are not available. Data sharing in the public sector is ordered by the 2012 Law on the reduction of administrative burden. The above-referred Law obliges all institutions to share their data so that to avoid double reporting. In Luxembourg, Article 12 of the 2011 Law specifies that the NSI should favour the use of administrative files. Surveys or censuses should be undertaken if the use of administrative data files is impossible or is not likely to provide reliable statistical information.

In Denmark, the Statistics Act also specifies that if the collection and processing of statistical information by a public authority or institution is being contemplated, then the NSI should be informed in order to negotiate potential coordination with the data owner. The data owners are obliged to inform the NSI on potential changes in their databases that may have impact on the final statistical output.

In Latvia a new Statistics Law, containing several provisions to ensure compliance with the EC Regulation No 223/2009 came into force in 2016. The new Law obliges administrative data holders to cooperate with the NSI in creating their registers, databases and information systems in order to facilitate the use of administrative data for statistical purposes.

In April of 2017, the Digital Economy Act was passed into UK law, making significant amendments to the Statistics and Registration Service Act 2007, bringing significant changes, including granting the power to all UK public authorities to voluntarily share the information they hold with the Statistics Board, overcoming any other legislative block or obligation of confidence (albeit in accordance with data protection legislation). Moreover, the new legislation provides the Statistics Board with a right of access to information held by UK Crown bodies, with the provision that failure to comply may result in the relevant correspondence being laid before Parliament. The Statistics Board is also granted a right of access to information held by all non-Crown public authorities, while failure to comply is a criminal offence. As well as requiring disclosure of information the notice may also require the public authority to consult the Statistics Board in advance of making any changes to its data processing systems. All such data shares must still be in accordance with data protection legislation. As regards access to private data owners, the Statistics Board is granted the right to access information held, whereas failure to comply with the corresponding notice is a criminal offence. As well as requiring disclosure of information the notice may also require the undertaking to consult the Statistics Board in advance of making any changes to its data processing systems. All such data shares must always be in accordance with data protection legislation.

2.1.2. Laws governing data sources

NSI's may also directly lobby, exert influence or actively participate in formal deliberations towards favourable revisions of legal acts governing specific data sources. In this case, provisions for allowing access to NSI may be included in the specific data source specific legal acts. For example, the Polish NSI actively tries to influence legislation in this direction. In the Czech Republic, from 2014 onwards, other state authorities are requested to explicitly state in their legislative proposals whether a new administrative data source is going to be established or whether an existing one is planned to be modified.

Finally, in Latvia, a national strategy on data sharing in the public sector is prescribed in the Law on State Information Systems of 5th June 2014.

2.2. Institutional and Organisational Aspects

Strong legal frameworks are a necessary condition for access to administrative data but not a sufficient one. They require adequate operationalisation in practice. In that sense, in conjunction with a general commitment to the exploitation of administrative data sources, the establishment of a visible institutional presence for administrative data within NSIs matters, as do more general organisational arrangements that raise the profile of the underlying activities and promote coordination across different actors.

2.2.1. Internal coordination practices

When production of official statistics is characterised by stove-piping, with organisational silos in charge of the entire life-cycle of a statistical production process, the possibilities for efficiency through the use of similar input data, the transfer of know-how, and the cross-fertilisation of good practices is frequently hampered. According to peer reviewers' assessments a formal, even statutory, coordinating role within a NSI, such as a dedicated coordination unit, would result in enhanced visibility of the administrative data and would increase efficiency through their common use within the NSI by avoiding duplication of effort, use of common methods, tools and software.

A dedicated unit (or key actor) may be assigned within the NSI with responsibilities for the promotion and use of administrative data, as well as for contacts and communication with the administrative data owners. In Belgium, the setup of dedicated services within the NSI to follow-up the implementation process for specific projects has been identified as a success factor. In Ireland, an Administrative Data Centre (ADC) has been set up as a special unit, in charge of developing and coordinating the statistical use of administrative data within the NSI and with administrative data authorities. The ADC acts as the focal point in receiving the administrative data, and the identification of potential new administrative data sources. It is also responsible for setting the quality and technical standards for the statistical uses of administrative data.

In Denmark, data from administrative registers of public organisations are centrally edited and organised by the NSI into statistical registers, which are highly harmonised across subject areas and over time. The assignment of unique identifiers for persons, businesses and buildings allows the intuitive linking of information among the various registers used. This results in only a small proportion of the statistical data being collected through surveys on businesses or households.

In Cyprus, the government operates a Government Data Warehouse (GDW) which is a central repository of data that enables easy access to accurate, consistent and integrated government data. The GDW is a single cohesive database with a subject-centric approach, optimised for reporting and analysis. The database contains selective transactions and inter-related information from all Government Information Systems, specifically structured for dynamic queries and analytics. The NSI has access to the GDW and, within the NSI, the Methodology Unit is responsible to coordinate the NSI's contacts and cooperation with the administrators of the GDW.

Within the Italian NSI, a Directorate was created, responsible for data collection of both survey and administrative data. This Data Collection Directorate is also in charge of data integration with the aim of supporting statistics production from registers. Its task is to acquire, store, integrate, and evaluate the quality of administrative data and make them and their metadata ready for use for statistical purposes. Domain experts do not belong to this centralised unit but are coordinated centrally.

This includes a massive documentation system on the metadata of administrative archives, the Documenting Public Administration Archives (DARCAP) system. The DARCAP system includes three main subsystems, namely: (a) DARCAP-DOCUMENTA used for the management of documentation on content and quality of administrative data sources, obtained by an analysis carried out jointly by the NSI and the administrative data owners; (b) DARCAP-INNOVA which collects information on changes in the administrative sources, and; (c) DARCAP-CONSULTAZIONE, which contains the collected information about the available administrative databases, their content and a general assessment on their quality. The information recorded in the DARCAP system is accessible by all members of the Italian Statistical System.

At Statistics Estonia, a specific unit within the NSI is in charge of coordinating contacts and cooperation with the data owners, while the Methodology Department and the Data Processing Systems Department are responsible for all administrative data collection. The Methodology Department negotiates with the owners of administrative and other registers, and concludes agreements with them. It is also in charge of the description of data in a central metadata system. Technically, the data from registers are transmitted to the NSI via a single-entry point which is under the responsibility of the Data Processing Systems Department. The latter is responsible for pre-processing the data and making them available to NSI's in-house applications. Through this procedure, it is ensured that there is no duplicated data and that the data are ready for statistical analysis. Additionally, in 2011, the office started developing the System of Statistical Registers (SRS) which integrates existing statistical registers (economic entities and agricultural holdings) and new statistical registers (persons, buildings and dwellings) into a common system. A new data collection system, called ADAM, is used for storing data in raw databases and for making data available to NSI's in-house applications. It also permits the extraction of detailed personalised information from administrative sources using X-road (national data exchange layer), ftp, etc.

Statistics Finland's Havas project (implemented from 2013 to 2015) aimed to identify and record all the administrative data collected and used by the NSI, build a centralised IT system for receiving administrative data, build a SAS system for pre-controlling the data received, describe all the data and variables in a centralised metadata system, and organise a team in charge of centralised administrative data collection. The team now handles about 65% of all administrative data files via this centralised platform, which utilises process metadata to control and inspect the received administrative data. All data are initially received on a Transfer Data Server, which is outside Statistics Finland's own data network.

In Hungary, a new and integrated data transmission system, called KARAT, has been developed. Through this system, the NSI receives a number of datasets from secondary sources, mostly from administrative sources. In parallel with the development of the KARAT system, the NSI redesigned their data source register used to manage the secondary sources. Currently, the data source register contains information on the administrative and other sources used in the data collection process and records information with regard to the data transmitted from other organisations. Furthermore, to support the NSI's transition to register-based statistics, a collection system for administrative data has been developed by the IT Centre of the Ministry of Finance.

In the UK, plans for the future include the organisation and storing of all data centrally, regardless of the source they stem from (administrative or survey data).

In general, there is an emerging need for training of the staff on the procedures related to the use of administrative data. In this respect, it has been suggested by the representatives of the NSIs during

the workshop organised by this project on access to administrative data sources (2016) that handbooks dedicated to such activities be produced. Moreover, efforts may be needed for a cultural shift so that staff are persuaded that use of administrative data will not hamper their work. Seminars or workshops are an important means for such discussions and exchange of opinions, possibly also with staff from other NSIs for the sharing of experiences.

2.2.2. Cooperation with other stakeholders

Broader coalitions resulting in wider public sector collaborations are also very conducive to the expansion in the use of administrative data, as well as the establishment of good practices.

In France, the Quality-Label Committee of the French National Council for Statistical Information (CNIS) rejects any proposals for new surveys when existing administrative data can be used instead. In a similar way, the Lithuanian NSI's policy prescribes that data from respondents should be collected only if administrative or other statistical data are not available.

In Switzerland, the NSI has adopted standard procedures to monitor developments in the legal framework governing the sharing of administrative data, as well as to identify administrative data sources potentially useful for statistical purposes. A register strategy has been established, with detailed records on which administrative data sources have the potential to be useful in the compilation of statistics.

Apart from close cooperation with data owners and providers for matters of access and use, cooperation and good communications with a variety of other stakeholders associated to varying degrees with the statistical process is desirable. This will ensure that they are aware of the importance of accessing administrative data for statistical purposes, of the associated benefits derived, as well as of the specific needs of statistical production.

Prominent among such stakeholders are often national data protection authorities. In Portugal, close cooperation with the data protection authority at an institutional level proved to be an important success factor for obtaining access to the required data sources.

Other stakeholders may assist in providing know-how, such as in the case of Portugal again where the Chamber of Chartered Accountants had an important contribution to the implementation of an IT system for access to business data.

Cooperation of NSIs with data owners and other stakeholders does not have to be restricted to specific data access cases. Cooperation networks bringing together NSIs and data owners help advance common strategies, standards, the exchange of experiences, and the establishment of good practices.

In Ireland such a formal network involving all producers of European statistics, the Formal Statistician Liaison Group (FSLG), has been set up to communicate and discuss issues of common concern. The FSLG comprises official statisticians as well as data analysts and researchers, all of whom are involved in the compilation and dissemination of official statistics in government departments. The statisticians engage in or facilitate formal communications between the NSI and statistical units in the public service, monitor the development and implementation of the data strategies, provide advice, and promote adherence to international principles.

In Lithuania, the Statistical Council advising the NSI on key strategy issues, which include the use of administrative registers for the production of official statistics, includes representatives of state and municipal institutions and agencies managing statistics, universities, public organisations, organisations representing businesses and the media. In the UK, employees of ministries on secondment to the NSIs act as mediators between NSIs and ministries (data owners). In Portugal, the NSI has been actively involved in the Simplex programme (administrative modernisation strategy - simplification programme for central and local government).

In Norway, the NSI established the Statistics Council with the view to strengthen the coordination among the producers of statistics. This consultancy body meets once a year and consists of the 25 major producers of statistics or central register owners. Its role is to coordinate the production and dissemination of official statistics and to prepare common rules and practices regarding the quality assessment of the data and the protection of confidentiality.

The Slovenian NSI is consulted and represented in discussions when any changes in administrative procedures that may have an impact on source data are to be implemented. The NSI seeks to influence the design of administrative databases through several means, including through governmental coordination procedures, a special Statistical Advisory Committee on Administrative Data Sources, and bilateral agreements and protocols signed with the administrative data providers. Moreover, in Slovenia, the main task of the Statistical Advisory Committee on administrative data sources is to monitor the developments in the legal framework and promote the exchange of experiences, good practices and information among organisations. Domain-specific information is exchanged among subject-matter statistical advisory committees.

In Ireland, the National Statistics Board (NSB), established by the Statistics Act, provides support and guides the NSI by establishing priorities for the development and compilation of official statistics. Among its activities is the power to arbitrate between the NSI and the other public authorities on the potential administrative data sources that can be used for statistical purposes where any uncertainty exists. In general, the NSB plays an active role in the promotion of the exploitation of administrative data. Moreover, the NSI has been actively promoting the use of an integrated approach for defining common identifiers, classifications and geo-spatial/postcodes across all administrative data sources. This requirement is emphasised in the NSI's statements of strategy. The NSI aims at providing technical assistance and support to Central Government Departments on the development of the national data infrastructure, advocating for a wider use of Personal Public Service Numbers (PPSNs), postcodes and promoting the usage of unique business identifiers across all relevant administrative records. Also, the Irish Public Service Reform Plan requires the NSI to develop a code of practice and standards for the gathering and use of data for statistical purposes in the Irish Public Service. Only those statistics assessed as compliant with the Irish Statistical System Code of Practice (ISSCoP) can be designated as official statistics and published under the ISSCoP logo.

2.2.3. Communication with the public

Public perceptions of the use of personal information by administrations or concerns about privacy and protection of sensitive personal data may prove to be an adverse factor if such concerns are not alleviated. On the other hand, public pressure may also prove to be a favourable factor for the expanded use of administrative data.

In Lithuania, for instance, the constant pressure exerted by enterprises on the NSI to use data already provided to the tax authorities was an important factor to the success of the usage of Value Added Tax data for statistical purposes, while Slovakia experienced similar pressures by enterprises to reduce response burden. Public trust is a crucial factor too. For example, in the Netherlands, the public trusts NSI's integrity and expertise.

As stated by the UK's ONS representative in the workshop organised by this project on access to administrative data sources (2016), the ONS has created a special media team to communicate what the ONS uses administrative data for.

Public trust is one of the fundamental principles of the NSI in Ireland as well. A stated policy principle of the NSI is "to be trusted by individuals, households and businesses to keep their data confidential, safe and secure and to operate in a way that is both ethical and legal (in both the spirit of, and to the letter of the law)". The NSI is publicly committed to the principle of good governance, emphasising data confidentiality and data security, creating a specific mind-set for conducting its business.

2.3. Cooperation with Data Owners and Wider Institutional Environment

The weight of legal frameworks notwithstanding, it is inconceivable to discuss success stories or perceive of good practices in the realm of administrative data in the absence of good cooperation between NSIs and data owners. While acts that provide strong mandates for access to the data and policies that encourage the sharing of data between public bodies are vital, the fact remains that these data were collected, processed and maintained for other-than-statistical purposes. Their useful transformation for purposes beyond their primary goal depends crucially on such cooperation.

2.3.1. Establishing the cooperation

Good practices involve substantive communications and the establishment of detailed processes covering protocols for the transfer of or access to the data, frequencies and timelines, metadata, contingency plans and many more. Such elements are also useful in communicating to the data user communities and the general public, informing them of the benefits. These manifest themselves in real gains for all, such as reduction of data collection costs and efforts, improved timeliness and quality. In Portugal, this factor has been identified as a major success factor for specific access cases and the NSI devotes considerable effort for communicating to the data owners the importance of obtaining the data for the compilation of statistics.

Successful cooperation between NSIs and administrative data providers is best seen as a mutually beneficial partnership toward a common goal rather than a merely dry administrative process. In that vein, a good practice is that personnel from the data provider are also engaged in the process in order to achieve actual commitment and goodwill at all hierarchical levels of the organisations involved.

Then, in the spirit of cooperation, it must be understood that the process incurs real, tangible costs and overheads for the data owners (in terms of human resources and time), which can prove significant if modifications and enhancements of their own processes for data collection, processing and maintenance are required. Therefore, gains for the data provider should be identified and communicated in good faith. These involve improvements in data quality which may prove beneficial for the data owner, potential benefits from the statistical processing of their data for their own management and decision-making needs, and gains from the statistical know-how of the NSI. This way, sustainable partnerships grounded on win-win scenarios emerge rather than a seemingly unilateral transfer of data from the owner to the NSI. Thus in the Netherlands the CBS had produced brochures for starting businesses, in which basic statistics for certain industries were presented. These brochures were distributed by the Chambers of Commerce, which is the most important basis for the CBS statistical business register and where all businesses must register.

Once the engagement and commitment of the data provider are secured, or at least encouraged, procedures for actual cooperation at different levels become necessary. From the findings of the survey, the need for close cooperation in three distinct areas emerged: (a) handling and overcoming potential legal obstacles or restrictions concerning the confidentiality of data and the protection of personal privacy; (b) potential difficulties concerning the suitability of data for statistical purposes or other issues having impact to statistical quality, that need to be overcome and may require suitable actions from the part of the data provider or modifications and enhancement of the register/data collection; (c) cooperation in setting up a frictionless and timely delivery of (or access to) the data, which extends to the actual technical requirements concerning data formats, technological infrastructures etc. These activities need to be coordinated at the strategic level, with clear schedules and aims.

Therefore, liaison mechanisms are necessary for these three parallel activities. These may take different forms ranging from common groups or committees to a variety of other liaison roles.

Besides an agreement for the provision of data, relevant modalities and a permanent cooperation mechanism can be established and documented for two primary aims: (a) effective cooperation on

actual data quality; and (b) effective cooperation in the cases where modifications to the register/data collection are required from either side.

The suggestions that follow emerge from the in-depth analysis of the problems reported by the NSIs in their national questionnaires, the insight obtained from the overall information collected (see Section 1.1) and relevant documentation identified on best practices.⁹

Effective cooperation on data quality involves the application of common, known and documented data validation rules and other quality checks. Ideally these rules and checks will be implemented as upstream as possible in the data flow, i.e. encapsulated in the data collection methods and tools of the data provider or in their own quality procedures before the actual delivery of the data. Moreover, the NSI should be able to send their own validation checks to the data provider and request corrective actions.

Modifications in the structure of registers/data collections affecting statistical concepts, classifications, coverage etc., may be required at any point in time and be initiated from either side. In any case, modifications from the part of the data provider that affect the statistical usage of the data must be communicated well in advance to the NSI and agreed on at the level of actual implementation. This requires a clear and documented procedure which has to be respected and efficiently implemented from both sides, with assigned roles and responsibilities. The procedure should describe in a clear and unambiguous way which modifications need to be communicated in advance, to whom and when, as well as conflict resolution mechanisms.

The bilateral agreement between the NSI and the data provider should also extend to metadata.

Apart from the provisions of formal written agreements, and particularly with an emphasis on the quality dimension, several good practices of cooperation have been identified. For example:

Statistics Norway regularly applies quality checks to the main administrative sources, following a set of quality indicators. Quality reports are being drafted by the NSI and consulted by the register owners.

In Germany, since 2016, the NSI in cooperation with the SLOs started to develop a strategy for the provision of tools and rules for the systematic examination of the quality of the data collected by the administrative data owners. Moreover, the Federal Statistical Office and the Statistical Offices of the Länder seek to participate in the so-called advisory boards of users set up by owners of administrative databases in order to be consulted on the design and make-up of administrative databases.

In Poland, representatives of the NSI actively participate in the working groups of the Ministry of Environment and other institutions and are in direct contact with the personnel of the ministry for clarifying issues related to the quality of the data.

In the Netherlands, the NSI has invested heavily in establishing a good cooperation framework with the data owners in order to improve the quality of the data by working together on the development of common standards and classifications.

In Latvia, domain experts are assigned to discussion with the administrative data owners for each and every specific issue that may arise.

In Ireland, a High Level Liaison Group is set up with each administrative data source owner and is convened at least twice per year to discuss the cooperation with the CSO and deals with any issues of concern that arise from either party in a transparent and open way. Moreover, the NSI has been actively promoting the use of an integrated approach for defining common identifiers, classifications and geo-spatial / postcodes across all administrative data sources. This requirement is emphasised in the NSI's statements of strategy. Implementing the institutional strategy 2015-2017, the NSI works on

⁹ United Nations, "Using administrative data and secondary sources for official statistics, A handbook of principles and practices" (2011).

the provision of technical assistance and support to Central Government Departments on the development of the national data infrastructure, advocating for a wider use of the Personal Public Service Number (PPSN) and a postcode and promoting the usage of a unique business identifier across all relevant administrative records.

2.3.2. National Data Sharing Strategies

Beyond legal acts per se, the existence and effective implementation of a national strategy for the sharing of data among administrations in the public sector has been identified as a success factor in several countries (e.g. Belgium, Spain). This is especially true when NSIs cooperate with government agencies tasked to oversee and coordinate these strategies (Belgium). The rationale underlying such strategies is often broader than the mere reduction of response burden by substituting administrative data for surveys. The guiding principle is “collect once and use multiple times as needed”, in a way that individual citizens and businesses do not have to provide multiple times information already available to one authority (Spain).

A supporting culture through strong political commitment plays a key role in specific legislative initiatives. Included here are the articulation of rules for data sharing among public administrations (e.g. Ireland), the set-up of general rules of cooperation between the owners of administrative data and national and local governments (Hungary), and the legal requirement for authorities to state in their legislative proposals whether a new administrative data source will be established or whether an existing one is planned to be modified (Czech Republic).

In addition to legal provisions, prevailing norms are also crucial. It is good practice for NSIs, for example, to participate proactively in the public discourse and the deliberations preceding any relevant legislative acts to ensure that statistical needs are taken into account and that they will not be hampered by any requirements imposed on the access to and use of administrative data.

Besides formal endorsement in the form of national strategies and legal frameworks, the actual strong political support from the part of government actors may prove useful in overcoming the reluctance of data owners. This may take the form of actual championship of a specific project at a ministerial level. For instance, in Portugal, the active support of the Justice Secretary of State, the Ministry of Finance and the Ministry of Presidency was crucial in the successful development of an innovative system, a central electronic system for collecting data from businesses (the Simplified Business Information - IES).

Purpose-driven practices are also effective, such as the set-up of specific government agencies or task forces to oversee and coordinate large scale projects for administrative simplification across the entire public sector. In these cases, the NSI's close cooperation with the respective agency contributes towards an increased and more effective use of administrative data. This has been identified as a success factor in the case of Belgium, where the close cooperation of the NSI with the corresponding agency (i.e. the Ministry of Finance) accounts for the Belcotax data source has been reported as an exemplary data access success case.

Other good practices by NSIs concern proactive communications at the governmental and public administration levels, whether formal or informal. In Austria, among the main tasks of the Statistics Council (the advisory board to the NSI) is the issuing of recommendations on the design of administrative data used for statistical purposes and the communication of these recommendations to the responsible ministries. In Estonia, the Director General of the NSI is required to submit to the Ministry of Finance an annual report on the implementation of the NSI's statistical programme, which must include a dedicated section on the quality of administrative records.

National strategies for data sharing may extend to the private sector as well. For instance, in the Netherlands, the NSI has the right to acquire data from registers maintained by any legal person (i.e. including private entities) when these registers have resulted from activities wholly or partially financed by the state. The new French Digital law adopted in 2016 makes it possible for the NSI and

the French Other National Authorities (ONA) to access to private data owned by private entities, for statistical needs. The new French Digital law asks administrations to share data between themselves and to promote open data as much as possible.

2.3.3. Interoperable national infrastructures

National initiatives and regulations to ensure interoperability between information systems in public administrations are also important in facilitating access to administrative data for statistical purposes. These include the set-up of appropriate infrastructures or mechanisms, such as the compilation of inventories of registers that include administrative data collections across the whole realm of public administration (Hungary). In Estonia, changes in the administrative databases are recorded centrally and are communicated to the NSI by the relevant authority. In Norway, a central hub is implemented where all administrative data of public authorities are deposited, while any public authority is forbidden to collect information that is already available in the hub directly from citizens, households or enterprises. Similarly, in Cyprus, a central government data warehouse exists where most of the administrative information is collected and maintained, and which is accessible by the NSI and all public authorities. In the UK and France there are policies for publishing government data in open formats, while in Poland there are ongoing efforts for technical standardisation that would enable interoperability among government databases.

2.3.4. Synergies with e-government applications and digital services

A promising approach is to target administrative data collection as upstream as possible, e.g. accessing the administrative data at the point of their actual collection (as compared to accessing compiled data sets in batch form). This is made possible by establishing interoperability mechanisms with e-government applications and electronic forms used by citizens and enterprises to complete administrative procedures.

Such successful cases have been reported in Belgium, while in Germany data from businesses are partly obtained directly from the information systems of businesses via special software modules voluntarily integrated in their accounting software.

In France, SBS data are derived from the ESANE system (Production of Annual Statistics on Enterprises), which combines information extracted from the incorporated industrial and commercial data on firms and from the DADS system (Annual Declarations of Social Data), which contains information on workforce size and wages.

Another good practice, tested in Bulgaria and proposed in Greece, is the implementation of single-entry e-government services, via which economic units can submit their data both to the NSI and to the National Revenue Agency, thus avoiding duplications and reducing significantly respondent burden. The data are stored in a common database, and each institution uses them for its own purposes, that is, the NSI for the compilation of SBS and annual accounts (balance sheet, revenue-income statement, etc.) and the NRA for fiscal purposes.

In the field of justice statistics, court information systems can provide similar data access opportunities. In Poland, the NSI accesses the data of the national court register via a publicly available web application, while in Germany the data collection form used in all Länder to register court data is also used to accommodate their simultaneous transmission to the NSIs. Thus, the form used by the courts to carry out their everyday business is also used to transmit statistical data directly to the Länder statistical offices in a transparent way.

2.3.5. Other Cooperation Aspects

Joint development of common IT tools: In Switzerland, the NSI has undertaken, jointly with Swiss Customs, the development of common IT tools for the processing and exchange of customs data.

Shared responsibility for joint production of statistics: in Iceland the agreement with the Central Bank covers, for example, the shared responsibilities for the production of Financial Accounts statistics and the supply of firm-level data for Balance of Payments statistics. In Slovakia mutual data exchange with administrative owners is agreed when feasible.

Recognition by the data owners of the importance of the data for the NSI: for instance Spain and Estonia identified as success factors contributing to a successful access to specific data sources the acknowledgment on behalf of the data owner of the importance of the data for covering the statistical needs of the NSI or data owners' understanding of the NSI's needs.

Provision of methodological support to the data owner: for example in Cyprus the NSI provided consultation on the classification of new entries to the register, mapping of data from NACE Rev1.1 to NACE Rev.2). Also, CYSTAT cooperated closely with the Social Insurance Services for the transition of the Social Insurance Register (SIR) from NACE Rev.1.1 to NACE Rev.2. The change was successfully implemented thus making the SIR data ready for use for statistical purposes.

Standardising written agreements: in 2015, in Finland, the NSI has started revising its data sharing agreements with the data providers to include specific provisions on the timeframe that the data owners should notify the NSI on planned changes to be implemented in the data source. In Malta the NSI has come up with the need to further standardise existent MoUs with the data owners in order to ensure that all necessary requirements to safeguard data quality are documented. Written agreements on the workload assigned to both organisations are signed with the data owners in cases where there are conflicts with their ambitions in Lithuania.

2.4. Quality and Technical Aspects

2.4.1. Methodological Work

In several cases (e.g. Austria, Finland), peer reviewers pointed out that dependence on administrative data makes the production of statistics vulnerable to changes implemented to their structure or content. This leads to the more general issue of data quality. Whereas historically NSIs have highly developed guidelines, standards and metrics for data quality related to traditional statistics, quality is either not easy to ascertain when it comes to administrative data or is subject to different criteria. Therefore, as NSIs expand more and more into administrative data they are aware of such matters. Good practices have already started to emerge, though.

In Italy, the NSI initiated research focussing on assessing the quality of administrative data and comparing them with survey data. Moreover the NSI has developed a standardised Quality Report Card for administrative data (QRCA), which provides a practical overview of the relevant quality indicators for a given dataset. The Integrated System of Microdata Repository (SIM) has been established as a subsystem of the overall metadata system and quality guidelines for use of administrative data have been prepared and published.

In Lithuania, the Methodology and Quality division of the NSI established a working group, responsible for the proposal of standardised methods that could be applied for editing and imputing survey and administrative data. In Slovenia, the NSI studied how the standard concepts of the quality assessment framework can be adjusted for the purposes of the quality assessment of the registers and large administrative data sets. The Slovenian business register was used as a pilot case¹⁰. In Sweden, the NSI has invested substantially in the development and use of new methodologies and quality assurance procedures with respect to its statistical processes and outputs, including the

¹⁰ Statistical Office of the Republic of Slovenia, Rudi Seljak, "Quality assessment of the registers and large administrative datasets" (2012).

integration of administrative data in the statistical production chain. In 2011, the NSI issued a publication on the quality assessment of administrative data¹¹.

In the Netherlands, the NSI has committed significant resources to the development of methodologies on the quality assurance of data derived from registers, which are used across the statistical domains. Recently, the NSI has also published a paper elaborating on a list of quality indicators that must be taken into consideration when checking the quality of the administrative data.^{12,13}

Eight NSIs are currently involved in the ESSnet on quality of multisource statistics¹⁴ (Denmark, Ireland, Italy, Lithuania, Hungary, Netherlands, Austria and Norway). The aims to develop and promote quality measures for evaluating the quality of administrative data and of the statistical outputs that use administrative sources in combination with other sources. The ESSnet will produce quality guidelines for the outputs based on multisource statistics.

2.4.2. Exploitation of technological standards

Information technology infrastructure that support centralised data collection from administrative data owners may facilitate more efficient and systematic processes, as well as provide better internal controls and transparency with regards to the administrative data used.

Additional good practices relate to the adoption of certain standards. Existing technological standards provide opportunities for transparent access to administrative data, when already used for other administrative purposes. “eXtensible Business Reporting Language” (XBRL) can be used to collect business data when businesses use the standard to report their financial data to authorities, while “Data Documentation Initiative” (DDI) and Statistical Data and Metadata Exchange (SDMX) can facilitate access to microdata or aggregate data already disseminated by administrations as open data.

In the UK, a pilot for the automatic extraction of SBS data directly from company files using XBRL has been implemented. The XBRL standard has been also adopted by the National Bank of Belgium and the Belgian Ministry of Finance in cooperation with the NSI. The use of a common standard allows the enterprises to automatically complete and transmit annual accounts, corporate taxes and statistical data to the three organisations.

At a more general level, a standardisation framework that supports the use of XML in the German public administration (XÖV) has been developed. The XÖV framework has been designed with the view to promote the systematic use of common IT standards for the electronic exchange of data in the federal public administration. The framework is used for the transmission of several administrative datasets, parts of which are used for statistical purposes.

In Italy, the necessity to design IT processes for the acquisition of administrative data has led to the creation of the AD Acquisition System, ARCAM. Through a special web site, the system allows ISTAT to acquire data in accordance with current legislation on data confidentiality and to preserve data integrity. This web interface can be a useful tool for communication with the administrative data suppliers in order to share the commitment to the use of data for statistical purposes and to improve usability/quality levels. The acquired data are stored in a single repository and are made available to internal users in a timely manner through the ARCAM management of permissions for data access. Administrative datasets that include microdata enter into the standardised process of the System of Integrated Microdata (SIM) and each administrative object, identified as a statistical unit (individuals,

¹¹ Statistics Sweden, Thomas Laitila, Anders Wallgren, Britt Wallgren, “Quality assessment of administrative data” (2011).

¹² Statistics Netherlands, Piet Daas, Saskia Ossen, Rachel Vis-Visschers and Judit Arends-Tóth, “Checklist for the quality evaluation of administrative data sources” (2009).

¹³ Statistics Netherlands, “Quality guidelines 2014, Statistics Netherlands’ quality assurance framework at process level” (2014).

¹⁴ https://ec.europa.eu/eurostat/cros/content/essnet-quality-multisource-statistics-komuso_en

economic units and places), is integrated. This data integration is performed through record linkage procedures ending with the assignment of unique codes that remain constant over time. Access to integrated data for internal users is also regulated in compliance with the law on data protection. The web interface also acts as a cooperation tool for effective communication with administrative data suppliers.

Estonia is moving to a system of machine-to-machine interchange of data, based on the XML format. The target is to achieve this by 2020. Data owners can apply for government funds to adapt their systems and make them compliant to that standard.

2.4.3. Extended Applications

Concerning the actual usage of administrative data, two application areas emerge as particularly promising, apart from the substitution of typical surveys by administrative data collections, namely pre-filling of electronic questionnaires and combining surveys with administrative data.

In several countries (e.g. Belgium, Czech Republic, Bulgaria, Germany, Denmark, Estonia, Italy, Netherlands) an effective combination of survey and administrative data is achieved. It includes a combination of numerous administrative sources that provide most of the required data, while the lacking variables, i.e. those not covered by administrative data, are collected as part of existing sample surveys, using estimation procedures to combine survey and administrative data. Examples include register based censuses, SBS, tourism statistics and agricultural statistics.

The use of administrative data for prefilling of questionnaires to reduce response burden has been successful in Belgium and Czech Republic. Similarly, in the agricultural sector, prefilled questionnaires based on information coming from administrative sources were used in the 2010 agricultural census survey in France, where respondents were asked to confirm or amend the information provided. In Lithuania, the e-Statistika system, developed by the NSI, is increasingly used by businesses for the provision of their data. The questionnaires included in the e-Statistika are prefilled with information from administrative sources.

3. Examples of data access cases

This section delves into administrative data matters as they pertain to specific statistical domains. It serves as a useful backdrop for the main findings contained in this report, while at the same time providing a perspective for their interpretation.

The analysis in this section is based on the NSIs' replies to the operational level questionnaires. The presentation relies heavily on the examples of data access cases selected to be reported by the NSIs. Thus, cross country and domain comparisons should be treated with care since the list of cases and issues is non-exhaustive and direct conclusions on the prevalence of such issues across domains cannot be drawn.

3.1. Data sources

An overview of representative “exemplary” and “problematic” data access cases, as reported by countries in the operational level questionnaire, are presented in Table 3.1. These cases reflect the actual experiences of NSIs as they practically attempt to exploit administrative data sources. Exemplary” and “problematic” cases have been selected to be representative of the general situation in each country concerning access to administrative data and to complement different obstacles and success factors in order to provide as a complete picture as possible (i.e. excluding peculiarities specific to a data source).

It must be noted that different data sources have been grouped under more generic headings, according to the administrative domain (e.g. taxation, population, social security, etc.), something that can conceal differences among sources under the same heading (such as vat, personal tax records, enterprise tax records etc.). This is unavoidable due to the differences in administrative and organisational structures among countries, which result in differences concerning data owners, organisation of registers etc. On the other hand, there is a considerable homogeneity concerning the reported target usages (i.e. target statistical domain) among sources under the same heading.

Interestingly, based on the number of cases selected and reported by the NSIs, countries can be easily grouped into three different sets:

- a) A group for which most countries had exemplary success (and a few encountered significant obstacles). This group is obviously the most promising and includes taxation, social security and population.
- b) An intermediate group for which some countries achieved exemplary success, namely business registers and property.
- c) A group where success cases are sparse and failures much more frequent, including energy, education, health, environment, agriculture and private data owners.

Nevertheless, misunderstandings due to potential misinterpretation of these results need to be avoided. A possible explanation is that cases (either successful or unsuccessful) for data sources in the first group are much more frequent because they are obvious candidates for data access. This may be so for a number of reasons that have to do with the specific administrative domains, such as high level of technological maturity and information systems required for the operation of taxation and social security systems.

Table 3.1. Data Sources Overview

● = Exemplary; ○ = Problematic

	Tax and VAT	Social Security	Population	Business registers	Property	Energy	Education	Health	Environment	Agriculture	Private data	Other
BE	●											
BG	○	●	●		○	●			○			
CZ	○	○	○		●							
EE			●	●	●					○		
IE	●										○	
EL	●○						○		○			○
ES	●	●										
HR	●	●							●	●		
IT	●	●		●		○				○		○
CY	○	●		●	●							
LV	●	●	●				○				○	
LT	●	●	●									
LU	●○	●				●						
MT	○											○
NL	●	●									○	
AT	●	●	●									
PL		●	○	○			●		●			○
PT			●○	●				○				○
RO			○	●				○				
SI			●							●		●
SK		●	●					●				
FI	●		●									●
SE	●		●		●							
CH	○●	●										●

Note: No information available for Denmark, Germany, France, Hungary, the United Kingdom, Norway, Iceland and Lichtenstein.

3.2. Success factors and obstacles

What follows is an analysis of the favourable factors and obstacles for data sources of the first group, where most countries had exemplary success. A similar analysis for the other two groups (i.e. the group where some countries only had exemplary success to report, and the “sparse” with occasional successes only) would be misleading due to the much more restricted number of cases.

In what concerns taxation, the most important factors reported were:

- High level of cooperation with data owners (10 out of 15 successful cases), while lack of an adequate level of cooperation was one of the main adverse factors (in 5 out of the 7 problematic cases);
- Existence of a written collaboration agreement (6 out of 15 cases);
- Favourable legal framework (4 out of 15 cases), while most of the problematic cases (5 out of 7) related to the legal framework (contradictions, restrictions etc.);
- The availability of data exchange technology was reported as a favourable condition in several cases (4 out of 15), while a lack thereof was identified as an obstacle in 3 out of the 7 problematic cases;
- Good data quality has been reported as one of the favourable factors in 4 successful cases, while, on the contrary, serious data quality issues were a problem for 4 of the unsuccessful ones;
- Excessive formality and bureaucracy were reported as a main obstacle in 4 out of 7 problematic cases.

While social security-related cases are fewer than those in the taxation domain, social security is the most successful domain, with 13 exemplary cases and only one problematic attempt. The most important factors reported were:

- Cooperation with the data owner as a success factor in all (13) cases;
- Other common favourable factors include the legal framework, data quality, availability of data, exchange technology (3 cases each).

With respect to the population domain, i.e. the exploitation of different available population registers, the main favourable factors included:

- High degree of cooperation with data owners (in 7 out of 10 cases);
- Favourable legal framework (5 out of 10 cases);
- Data quality (5 out of 10 cases);
- Availability of data exchange technology (5 out of 10 cases).

In contrast, in 2 of the 4 problematic cases it was the legal framework which created most obstacles, including room for misinterpretations and contradictions between the law granting the right to access and laws governing the data sources, restriction in the usage of the data for specific statistical purposes only or restrictions in the right to link data with those of other data sources or even the requirement for specific ad hoc legal acts.

Table 3.2. Taxation data sources – exemplary cases

Data source		Statistical domain	Success Factor
BE	Tax forms	Labour Cost Survey, LFS, SILC	Commitment of resources
IE	Employ level returns by registered employers	Business and household statistics	Collaboration agreement Cooperation with data owner
EL	Taxation data	SBS, STS, National Accounts, Business Register	Legal framework Collaboration agreement Data exchange technology Data protection technology

	Data source	Statistical domain	Success Factor
ES	Tax Authorities data/ personal income/ VAT / Corporation tax	SBS	Legal framework Collaboration agreement Cooperation with data owner
HR	Tax Authorities data / personal income / VAT / Corporation tax	Trade statistics, National Accounts, SBS, SBR, SILC, Labour Statistics	Collaboration agreement Catalogue of data exchange
IT	Tax data on individuals	SILC, Statistics on income, Labour Market statistics	Cooperation with data owner Timeliness of data release Data quality
LV	Taxpayers register	SILC, LFS, Population Census, Population statistics, Business Statistics	Collaboration agreement Cooperation with data owner
LT	VAT declarations	STS	Public opinion Data exchange technology Existence of uniform / unique identifiers
LU	VAT declarations	Business register, Business Statistics	Collaboration agreement Timeliness of data release
NL	VAT declarations	SBS, STS	Cooperation with data owner Data quality
NL	Income Tax Data	Statistics on income	Cooperation with data owner Data quality
AT	Tax Data	Business Register, Business Statistics, National Accounts	Legal framework Cooperation with data owner Data exchange technology
SE	Income Tax Data	Labour Market statistics	Cooperation with data owner Data quality Existence of uniform / unique identifiers
FI	Data on personal taxation	Statistics on income distribution, taxable incomes, population census	Legal framework Existence of uniform / unique identifiers Knowledge and experience on data content of the register
CH	Customs data (imports/exports)	Business statistics, Price statistics	Cooperation with data owner Data exchange technology

Table 3.3. Taxation data sources – problematic cases

	Data source	Statistical domain	Obstacle
BG	VAT declarations	STS	Lack of cooperation Inadequate coverage Incompatible definitions of concepts Data exchange technical issues
CZ	Income Tax Declarations (persons and enterprises)	SBS (directly), Business statistics (indirectly), Business register	Restricted to specific statistical purposes Excessive formality / bureaucracy Missing data items Data exchange technical issues
EL	Tax authorities data	SBS, STS, National Accounts, Business Register	Restricted to specific statistical purposes Incompatible definitions of concepts Missing data items

Data source		Statistical domain	Obstacle
			Missing records Data exchange technical issues
CY	Income Tax Declarations		Confidentiality restrictions Permission from Data Protection Authority required Restrictions on linking microdata Lack of cooperation Lack of methodological metadata Unsatisfactory Timeliness
LU	Tax Database	EU-SILC statistics	Legal framework contradictions Data owner refusal Confidentiality restrictions
MT	Tax Authorities data	SBS, National Accounts, STS, SILC, Gender Pay Gap	Requirement for ad hoc legal acts Legal framework contradictions Permission from Data Protection Authority required Restricted to specific statistical purposes Data owner refusal Non-respect of agreement Excessive formality / bureaucracy Missing data items Incompatible reference period
CH	Tax Authorities Data	Business statistics, Population statistics	Data owner refusal Excessive formality / bureaucracy

Table 3.4. Social Security data sources – exemplary cases

Data source		Statistical domain	Success Factor
BG	Register of insured persons	Labour Statistics	Legal framework Cooperation with data owner Timeliness
ES	Public Social Benefits Register	Population census	Collaboration agreement Cooperation with data owner Right granted by the data protection authority
IT	Social Security data		Cooperation with data owner Data quality Availability of metadata
CY	Register of employees - social contributions	Labour Market Statistics (SES, LCI, Employment statistics)	Cooperation with data owner
LV	Social Assistance Administration Information System	SILC, Population Census, Population Statistics	Collaboration agreement Cooperation with data owner Adoption of common technical solutions Data exchange technology
LT	Number of employees and wages and salaries	STS and Labour statistics	Cooperation with data owner NSI commitment Data exchange technology

	Data source	Statistical domain	Success Factor
LU	Employment database	LFS	Collaboration agreement Cooperation with data owner Data are not imposed to changes over time
NL	Data on Social Security and Income	Statistics on income, social security benefits and employment	Cooperation with data owner Data quality
AT	Central Social Security Register	Census, Labour Market Statistics, National Accounts, Business Register, SBS	Cooperation with data owner Data exchange technology
PL	Social Assistance in the place of residence	Living conditions, social assistance	Cooperation with data owner
SK	Policy holder databases	Demography, census, social statistics	Legal framework Cooperation with data owner Data quality
CH	Old-Age and survivors Insurance	Business, Social and education Statistics	Legal framework Cooperation with data owner

Table 3.5. Population data sources – exemplary cases

	Data source	Statistical domain	Success Factor
BG	Population Register	Demographic statistics	Collaboration Agreement Cooperation with data owner
EE	National Defence obligation register	Social and demographic statistics	Legal Framework Cooperation with data owner Data quality
LT	Data on persons	Demographic statistics	Cooperation with data owner Data Quality Data exchange technology
AT	Central Population Register	Population Statistics, Census, LFS, EU-SILC	Legal Framework Data exchange technology Cooperation with data owner
PT	Integrated system for civil registration and identification (SIRIC)	Demography, Population Statistics	Definition of statistical units and variables Access to metadata Data formats Cooperation with data owner
PT	Data Sources used for the feasibility study of 2021 census	Population Statistics	Cooperation with data owner Recognition by the data owners of the importance of the data for the NSI.
SI	Central Population Register	Demography, Census, Labour Market, Living conditions, Social and Health statistics	Experience with the data from the part of NSI Data quality
SK	Register of natural persons	Demography, census	Legal framework Good structure of the data
FI	Population information system	Population and vital statistics, survey frame, population census	Legal Framework Data Quality

Data source			Statistical domain	Success Factor
				Existence of unique identifiers and links between units
SE	National Registration	Population	Social statistics and especially population statistics	Cooperation with data owner Data quality Availability of linkage information
CH	Population Registers of Municipalities		Business, Social, population and education statistics	Legal Framework Data exchange technology

Table 3.6. Population data sources – problematic cases

Data source			Statistical domain	Obstacle
CZ	Register of inhabitants		Population statistics, Population census; Business register	Legal framework (room for misinterpretations in, contradictions, restrictions of right to use only for specific statistical purposes) Data owners' refusal Inadequate Coverage Incompatible data formats
PL	Databases of Registry Offices	Civil	Social Statistics, Demography	Excessive formality or bureaucracy inadequate completeness
PT	Civil Identification database (BDIC)		Population Statistics	Legal framework (misinterpretations, specific legal acts required, contradictions, Permission from data protection authority required, restrictions on linking microdata) Lack of effective cooperation mechanisms Incompatible definitions of statistical concepts Inadequate coverage Lack of unique identifiers
RO	Ministry of Internal Affairs		Population and Demographic Statistics	Lack of effective cooperation mechanisms (written agreements signed are not respected)

4. Concluding Remarks

Administrative data have been for a long time integral parts in the production of official statistics in some European countries. In recent years, there has been renewed interest in expanding their acquisition and use, both in breadth and in depth. Increasing demands for data on NSIs, coupled with budgetary constraints, have precipitated concerted efforts among all NSIs to capitalise on the promise of such data. Developments take place in a dynamic environment and through a multitude of parallel activities. Existing legal frameworks are revised or new ones are enacted, new alliances among public bodies are forged, public attitudes are changing, and a new data culture sets in.

Virtually every NSI is active on this front, most with a series of initiatives related to administrative data. This project is part of Eurostat's ongoing efforts to support progress in all areas related to the acquisition, curation and utilisation of administrative data sources by fostering dialogue and exchanges among NSIs. It involved desk research, literature reviews, customised surveys to European NSIs, as well as a physical gathering with participating NSI officials to discuss common issues, exchange experiences, and contemplate prospects.

This particular report highlighted numerous examples of good practice in a variety of areas. It also identified areas where impediments are encountered. While all these examples serve to illustrate points made rather than offer an exhaustive list, it is hoped that they can serve as useful references and contribute towards the continuous cross-fertilisation of ideas and practices among NSIs so that additional mileage is gained in the near future. This way, more of the promised benefits of administrative data will be materialised, generating efficiencies from time-consuming and expensive survey methods and limiting them to the absolutely necessary, reducing burden among citizens and businesses, and leveraging investments in a way conducive to smarter public administrations with logical and prudent "recycling" of already-collected data.