Title	**Audit report on access mechanisms and availability of official statistics across the European Research Area, with recommendation on role of the CESSDA Research Infrastructure in facilitating greater access (D10.1)**
Work Package	WP10
Authors	Paola Tubaro, Roxane Silberman, Marie Cros, Anne Cornilleau, Vigdis Kvalheim, Dag Kiberg, Peter Farago.
Source	WP10 team
Dissemination Level	PU (Public)

Executive summary

The CESSDA-PPP endeavours to upgrade CESSDA into a European Research Infrastructure Consortium for access to social science data, able to provide pathways for access to all or most of the sources of data that researchers need. This report focuses on government microdata, which constitute today a substantial part of the data needed for research and teaching in the social sciences. The broad category of government data includes surveys produced by National Statistical Institutes (NSIs) as well as administrative data and combined survey-administrative datasets. Today, social scientists' demand for governmental data has surged, with rising interest in particular for administrative data and for once peripheral fields such as health and the environment.

At present, access to these data is uneven across Europe and not systematically mediated by CESSDA members, while in recent years new challenges have arisen: improved statistical tools enable analyses based on detailed microdata, technological progress facilitates increasing use of administrative data bases and combined datasets, whereas at the same time privacy protection concerns are growing. Long neglected, researchers' needs are progressively gaining recognition in the legal framework, and governmental actors (including, but not limited to, NSIs) are setting up new forms of access both for anonymised and for confidential data. Driven by the European integration process, a parallel infrastructure led by NSIs and Eurostat for access to government data is increasingly likely to emerge, thereby challenging the role of CESSDA and undermining the very idea of a single, comprehensive European infrastructure for social science data.

This report aims to map current access arrangements throughout Europe, considering both the status quo and new or expected developments that may require attention in the years to come. It takes into account modes of access to both anonymised and confidential data, with special emphasis on the role of CESSDA member organisations as intermediaries. It considers both

the national level (NSIs and country-based institutions) and the European level (Eurostat data). It results from the work of Workpackage 10 in the CESSDA-PPP, which has also involved efforts to make first steps in order to propose solutions to existing problems, especially by starting talks with governmental actors. A milestone of this task has been the organisation of a joint Eurostat/ONS/CESSDA workshop on Microdata access in 2008.

The main results of the audit report can be summarised as follows:

- Overall, the legislative framework has evolved in a sense that is more favourable for research than in the past, and conditions for access have improved in many countries, despite remaining barriers.
- Many NSIs have recently been very pro-active in setting up new facilities for access, based on advanced technological solutions in order to protect confidentiality (particularly safe data laboratories, both on-site and remote through the internet).
- CESSDA organisations are currently very heterogeneous: while some members are substantially involved in the dissemination of government data and have accumulated significant experience in this area, others do so to a very limited extent, if at all.
- These developments reinforce fears that governmental data might in future be offered by a parallel, Eurostat and NSI-led infrastructure, outside of CESSDA.
- However, some CESSDA members are setting the example in establishing enhanced and renewed forms of cooperation with government actors (also including forms of access to confidential microdata).
- The output of the Eurostat/ONS/CESSDA workshop of December 2008 is also highly encouraging, suggesting a possible extension of existing cooperation schemes to CESSDA as a whole.

The main resulting recommendations can be summarised as follows:

Regarding CESSDA members:

- Significant involvement in government microdata dissemination should be a condition for membership of the cessda-ERIC;
- Thus, efforts to reach agreements with governmental data producers should be required of CESSDA organisations that do not currently offer any mediation services for government data;
- To achieve this, CESSDA should set up a permanent sub-committee, or expert group in charge of providing assistance to members that need to prepare a first agreement with governmental statistics in their home countries.

On the relationships of CESSDA towards NSIs and Eurostat:

- The cessda-ERIC should explicitly give a place in its statutes to Eurostat and possibly to other NSIs that are willing to cooperate.
- To do so, CESSDA should make an informed choice between different possible modes of operation.
- To inform its decision, CESSDA should in the short run set up a team or expert group to explore more closely the possible options.
- In addition, CESSDA should immediately start negotiations with Eurostat and NSIs on chapters where rapid progress is possible.

- CESSDA should promote the transformation of the Eurostat/ONS/CESSDA event of 2008 into a permanent forum (possibly taking place every two years).

On the relationships of CESSDA with the research community:

- Representatives of researchers should be members of the Scientific Council of the ERIC, and possibly of the sub-committee or expert group that is in charge of providing advice on partnerships with governmental statistics.

To achieve these goals, follow-up work has already been undertaken. Eurostat's willingness to continue and enhance the experience of the 2008 workshop facilitates progress in some areas. In particular, R. Silberman was invited to a Eurostat seminar on remote access in June to discuss possible forms of future collaboration, and the currently open FP7 call (30 July 2009) on Data Archives and remote access to official statistics is an opportunity for possible joint work by CESSDA and some NSIs with support from Eurostat.

1. Introduction

The CESSDA-PPP endeavours to upgrade CESSDA into a European Research Infrastructure Consortium for access to social science data. In this role, it is expected to provide pathways for access to all or most of the sources of data that researchers need. A substantial part of these data are government microdata, for which researchers' demand is high especially in the fields of economics and sociology. Some CESSDA members already have substantial experience of government data dissemination, but there are discrepancies across countries as others have not considered this area as their main focus so far. At the same time, the situation is rapidly evolving, with new challenges that will need to be addressed by the new ERIC.

Government data include surveys produced by National Statistical Institutes (henceforth, NSIs), many of which are of very high quality. Some of them are harmonised (Labour Force Surveys, LFS) or produced directly at European level (Survey of Income and Living Conditions, SILC), and are of great interest for comparative research. In some countries the statistical services of ministries, government departments, local authorities, and other public-sector agencies are also in charge of conducting surveys that are increasingly demanded by researchers. In addition, recent improvements in IT and statistical methods and tools have raised researchers' expectations for access to other types of government microdata, namely administrative data bases which provide valuable bases for longitudinal analysis and for public policy evaluation. Availability of these data is already wide in countries where registers have been used for long as an alternative to census; elsewhere, a tendency to replace surveys by administrative data and/or to merge survey and administrative data in order to reduce respondent burden, has substantially enriched these data sources and has made them even more attractive for social scientists. Newly emerging research needs and improved techniques to combine various datasets (governmental surveys and administrative databases; two or more administrative datasets; academic surveys and administrative data) are also building bridges between datasets that were once accessed through separate systems by different research communities. In particular, the frontiers between datasets that were traditionally used by social science researchers on the one hand and health and environment data on the other hand, are partly blurred. As a result, the set of government microdata of interest for the social sciences has substantially swollen and tends to inflate further, with new issues in term of access: in particular, data and documentation quality are not always up to

standard, and some of these data (especially tax and health data) are highly sensitive and often difficult to anonymise.

Encouraged by methodological developments and greater availability of new data sources, researchers increasingly demand detailed individual data, though in a context in which the media, public opinion, and many governmental institutions express growing concerns for privacy protection. Technological progress has contributed to addressing these conflicting needs with new modes of access that can better combine protection and researchers' need of detailed data. On one hand, statistical disclosure control techniques and web-based tabulation tools allow releasing datasets in highly anonymised versions; several NSIs have indeed reinforced their capacity to provide free and direct access to tables and anomymised Public Use Files (PUFs) through the Internet. This tendency tends to spread even though access to such data is still uneven across countries at the moment, with a few outlying cases in which it is almost inexistent. On the other hand, new forms of access to confidential data are being developed and implemented: Eurostat and several NSIs are in the process of setting up high-technology secure facilities for access to confidential data through on-site safe centres and/or secure remote connection systems. These two parallel developments are connected, to the extent that heavy anonymisation may jeopardize the usefulness of the information contained in the data and fail to meet researchers' expectations; this provides an incentive for researchers to shift their attention to confidential data. Further, the very perimeter of confidential data has enlarged as it now includes data that were once considered as sufficiently anonymised. These processes, in which NSIs and Eurostat are major actors, constitute new challenges for those CESSDA members that have so far played an important role in the dissemination of governmental statistics in their home countries.

Will these tendencies lead to the rise of a parallel European microdata infrastructure run by Eurostat and governmental authorities? This scenario must be taken very seriously as a potential challenge to the future role of CESSDA and the whole idea of a European infrastructure offering a single, uniform, and comprehensive system of access to social science data. One might ask, however, to what degree Eurostat and NSIs are able and willing to manage such a data infrastructure without participation of data archives and the research community. In fact, the present situation offers a unique opportunity for CESSDA to position itself at the very core of the European data system: indeed, there is evidence that Eurostat and a few NSIs (and other governmental institutions) would be willing to share with CESSDA some of the workload involved in providing access to data for research purposes, and would be glad to draw on its expertise and its close relationships with the research community, in order to improve their services while keeping costs under control. By engaging in an unprecedented effort to enhance its role as an intermediary for government data, CESSDA might reinforce its role while improving conditions for research throughout Europe.

The context in which a renewed CESSDA will have to act is nevertheless complex due to heterogeneities among countries at all levels: legislation, discrepancies between official procedures and actual practices, and what might be broadly called the culture, history and traditions of each country. Differences concern not only NSIs but also CESSDA members: they have traditionally operated at national level and do not necessarily share the same mission, perspective, and experience at the moment. The new ERIC will have to upgrade from the current situation to a stronger form of integration in which expertise will be pooled, coordinated and put into operation in a truly pan-European experience.

Within this framework, the report of Work Package 10, prepared by team members involved in tasks 10.1 and 10.2, aims to:

- draw an up-to-date picture of current arrangements for access to government microdata, also taking into account privacy protection concerns, often raised by NSIs and other data producers;
- describe the current role of CESSDA organisations in this context, with emphasis on similarities and differences between countries;
- identify main issues, trends, as well as remaining or newly created barriers;
- consider recent and new developments that may challenge the results already obtained by some members in improving access conditions through their intermediation;
- make recommendations about paths for cooperation and partnership between CESSDA, NSIs, other governmental departments or agencies, Eurostat, and the social science research community, in the perspective of upgrading CESSDA as a European Research Infrastructure Consortium.

The findings of the report support the claim that an upgrade of CESSDA calls for:

- reinforcement of existing agreements between CESSDA organisations on the one side, and NSIs and/or other statistical agencies or departments on the other side;
- preparation of new agreements in countries where they do not already exist;
- creation and maintenance of a Europe-wide formal framework for continuous cooperation and partnership, involving Eurostat, CESSDA, and national NSIs;
- involvement of researchers in the governance of CESSDA;
- immediate launch of negotiations in areas where CESSDA can make a contribution even in the short/medium run.

The remaining of this report is organised as follows. Section 2 details the Methodology used; Section 3 reports main findings, distinguishing between legal and institutional context, current access arrangements, costs, levels and providers of access, researchers' accreditation, and functions of CESSDA organisations. Section 4 presents the main conclusions and recommendations. Section 5 includes references.

2. Methodology

Information for this report has been collected from a variety of different sources in order to draw a picture of the current situation that could be as complete as possible. A major difficulty during the information collection phase has been the continuously changing nature of the European data landscape, due to modifications in the legislation of some countries as well as to an incessant evolution of practices. For this reason, regular updates of the information collected have been necessary, even within the short time span of the PPP. What follows is a brief description of the main sources of information that have been used and of the way this material has been exploited.

A first source of information is a survey that was distributed to all CESSDA members and to a small number of non-CESSDA data organisations. The survey included several sections with questions on various aspects of the data archiving activity. Some of these questions referred specifically to the role of CESSDA organisations as distributors of government data and more generally, to conditions of access to government data (including confidential data) in the country where respondents were operating. The CESSDA organisations that responded

to the survey are: ADP (Slovenia), ADPSS-Sociodata (Italy), CEPS/INSTEAD (Luxembourg), CIS (Spain), DDA (Denmark), EKKE-GSDB (Greece), FORS (Switzerland), FSD (Finland), GESIS (Germany), ISSDA (Ireland), NSD (Norway), RQ (France), RODA (Romania), SDA (Czech Republic), SND (Sweden), TARKI (Hungary), and UKDA (UK); DANS (Netherlands) completed it partially. The survey was released at the end of May and was closed in early July 2008[1].

Another source of information is a survey that was distributed to National Statistical Institutes of countries where CESSDA organisations are present, as well as to other government agencies or statistical services of ministries that are important data producers and providers in those countries. This survey has been preceded by a pre-survey distributed to directors of CESSDA organisations, asking them: 1) to name government agencies in their countries that could be invited to respond to the survey, together with the NSI; 2) to choose whether they were willing to have the Project Office invite their NSI (and possibly other agencies) to respond to the survey, or they preferred to contact them themselves. This procedure was meant to respect and preserve the relationships that some CESSDA organisations already have with public-sector producers of statistical data in their home countries. Almost all CESSDA organisations responded to the pre-survey, which made it possible to distribute the NSI survey to a large number of potential respondents. This was done at the end of August 2008, and responses were collected until the end of November 2008. Respondents to the survey include the National Statistical Institutes of Austria, Czech Republic, Greece, Finland, Hungary, Ireland, Italy, Romania, Slovenia, Spain, Sweden, and the UK, as well as a number of non-NSI government data producers, namely the DARES (the statistical service of the French Ministry of Labour); the Bank of Greece; the Employment Observatory Research of Greece; the Hellenic Migration Policy Institute; the Bank of Italy; the Norwegian Directorate of Immigration; the Norwegian Institute of Public Health; the National Board of Health and Welfare of Sweden; the National Council for Crime Prevention in Sweden[2].

A third source of information has been the collection of relevant literature, including published articles, working papers, reports, texts of laws, and web-based materials. A list of most of the material retrieved and used for the project is in section 5.

Some information has also been shared with the European Research network EQUALSOC. Indeed two WP10 team members are also in the Data Support Committee of EQUALSOC, whose mission is to facilitate members' access to national as well as to Eurostat microdata. To collect information for a report required by the European Commission, EQUALSOC research teams were asked to indicate which government data they were using and how they obtained them. This survey was conducted in autumn and winter 2008, thus offering an opportunity for synergy with the WP10 investigation.

To a lesser extent, information has also been retrieved from interviews and personal contacts with researchers and data users as well as with public-sector statisticians and staff members of CESSDA organisations. This has been done unsystematically, however, as a complement of information and not as a major source.

With this material, WP10 team members prepared intermediary country reports on access arrangements. Each team member was in charge of one or a few countries. These pre-reports were then used to assess similarities and differences across countries so as to obtain a more general picture of the situation. Some contradictions emerged among the different sources,

[1] The Data Archives questionnaire is available at http://www.cessda.org/project/doc/cessda_ppp_survey.pdf.
[2] The NSI questionnaire is available at http://www.cessda.org/project/doc/cessda_survey_nsi.pdf.

though, possibly reflecting discrepancies between official regulations and actual practices. Further clarifications may thus be called for in a follow-up report.

It needs to be added that WP10 was not expected to conduct its investigation in countries where no CESSDA members are currently present. Also, information has been collected unsystematically in some particular areas such as health data, unemployment and pension data, environmental data, geo-tagged data, fiscal and financial data. These data are not always part of governmental statistics (they are sometimes the by-product of the activity of specific administrations or of agencies with mixed private-public legal status) and access conditions vary widely depending on the structure of the statistical system and the status of the producer. For instance in Norway, administrative health data can be accessed through NSD, whereas in France a completely separate access system has been set up for social security datasets, with participation of various social and economic actors. This issue calls for careful watch in the future because an increasing number of social scientists work at the interface between these areas and more traditional ones.

3. Findings

This section presents, after a brief presentation of the legal and institutional framework for access, our main findings concerning current access arrangements, costs of access, the level of access and the type of providers, researcher accreditation issues with special emphasis on the problem of training future researchers and cross-border data circulation, and finally the functions of CESSDA organisations including not only dissemination and accreditation, but also various forms of user support.

3.1 Legal and institutional context

This sub-section briefly illustrates elements of the legal and institutional context that may have important repercussions on access.

3.1.1 Statistical Laws and Privacy Protection Laws

The legal framework that defines the possibilities and the limits for microdata dissemination to researchers consists, in most countries, of a Statistical Law together with a Privacy Protection Law. While the confidentiality safeguards of Privacy protection laws are meant to protect the fundamental right of every individual to privacy[3], those contained in the Statistical laws (or other relevant regulation, e.g. the national constitution in the case of Germany) aim specifically to maintain the trust of respondents, who may otherwise be unwilling to continue to cooperate in the data collection activities of public-sector statistical agencies. In some cases, these laws also attribute a specific status to national statisticians, e.g. by submitting them to professional secrecy obligations. In this perspective, a tension appears between confidentiality protection and the need of social science researchers to have access to sufficiently detailed data for fine analysis, model building, and policy evaluation.

The data needs of researchers were rarely taken into account when national statistical laws were originally drafted (mostly following World War 2) and when privacy protection laws

[3] Today, these provisions apply to any database (including government data) whose creation required even minimal direct or indirect individual identification information.

first appeared (mostly in the seventies). The first specific provisions in some privacy protection laws were introduced to meet the needs of epidemiologists (as in France in 1994). The 1995 European Directive on protection of personal data[4] recognized the research purpose and was followed by analogous developments at national level, in several countries (although pioneering efforts had already been undertaken, e.g. for access to business data in France). At EU level, the Commission Regulation n. 831/2002 addresses the question of access to confidential data for research purposes, and the Statistical Code of Practice adopted in 2005 includes provisions on the use of statistical data by researchers. Because EU regulations only concern the statistical activities that are carried out for the production of European-level statistics, specific confidentiality regulations subsist at country level, and are sometimes at variance with the EU statistical confidentiality regime, with substantial heterogeneities across Europe. Interestingly, however, a number of countries such as France (2008), Portugal (2008), and the UK (2007, in force from April 2008) have revised their laws so as to explicitly recognize researchers' needs, and to set up rules and procedures under which access to government microdata, including confidential data, is open to accredited researchers. For some countries, changes were even more sizeable due to the presence of statistical confidentiality principles in the national constitution (Germany). Further improvement is expected at EU level, with the new Commission Regulation on European Statistics (223/2009) now in force, and possibly at national level if other countries follow.

Other parts of the legal system need to be taken into account. In some countries, the National Archives store NSI data and give access to them, sometimes in confidential versions, especially to historians. In these cases, the legal provisions regulating access to documents held in the National Archives interfere with Statistical laws as they define conditions for access to confidential data, including NSI surveys and censuses as well as other statistical datasets[5]. In addition, some government data fall under specific provisions, e.g. administrative data (CADA law in France), fiscal data, financial and banking data collected by Central Banks. On the whole, the legal framework defining conditions of access to government microdata is complex and sometimes entails contradictions between different legal provisions in the same country. This complexity reinforces heterogeneities between countries in Europe.

Nonetheless, the needs of researchers tend to be taken into account in a growing number of countries, though at a different pace and under different conditions (including the set-up of high-technology solutions such as safe centres or remote access facilities). Also, Freedom of Information provisions, which have been introduced at European as well as at national levels, have further contributed to creating a more research-friendly environment.

In sum despite many remaining restrictions, the legal framework is progressively evolving in a direction that is more favourable to research than it used to be.

3.1.2 The structure of the national statistical system

Microdata release to researchers is also highly dependent on the structure of national statistical systems. In particular, a key factor is the degree of centralisation or decentralisation of statistical systems, both in a geographical sense (the possible role of regional or local statistical offices with respect to a central agency, which is important especially in the case of Federations) and in a functional sense (the possible attribution of important data collection

[4] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. *Official Journal,* L 281, 23/11/1995, pp. 0031 – 0050.

[5] Recent changes of the French Statistical Law were indeed introduced in connection with Archives regulation.

functions to the statistical services of ministries, government departments, and other public-sector institutions). Even in decentralised systems, data users may face one central provider or several providers, depending on whether a central agency coordinates the statistical activities of all actors. In some countries, lack of coordination is due to the fact that the statistical law does not apply to administrative data; in particular, fiscal data are often subject to specific rules outside the control of the NSI. This case is sometimes more favourable to researchers' access, in that decentralised systems often allow for more flexibility; in other cases, the opposite is true. More generally, the existence of several unconnected providers may be unfavourable to researchers as it increases the amount of time needed to search for appropriate data sources, to collect information and documentation, and to prepare application packages. For Data Archives, it likely increases the time and effort devoted to prepare and conclude agreements with data providers.

Another potentially important factor is reliance on register-based data collection rather than questionnaires; yet the traditional distinction between countries that mainly use registers, notably the Nordic countries and those that do not, is being blurred by recent efforts to reduce the burden generated by surveys. Together with a generalized increase in the number of available administrative sources, this tendency has recently led to a large-scale shift from survey-based to register-based statistics in many countries. In the near future, this may well enlarge the perimeter of governmental statistics of interest for social science researchers.

Other factors that need to be taken into account at least in some statistical systems are the possible role of Central Banks as data producers and/or providers. Differences in status and roles of specific government or mixed status agencies for particular areas of interest such as health, unemployment, and pensions, should also be taken into account as noted above.

Similar considerations apply to the European Statistical System. Researchers and Data Archives currently negotiate with their NSI at home and with Eurostat at European level. A tendency towards greater centralisation in the data production process appears, in particular in the cases of the European Household panel, SILC and LFS. Access is still heterogeneous at present: for some countries, only the Eurostat datasets are available, while others also make more detailed national samples available. Depending on the future evolution of the European Statistical System and its degree of centralisation, the respective role of NSIs and Eurostat as potential interlocutors for researchers, Data Archives and CESSDA as a whole may evolve.

The degree of centralisation of statistical systems, the use of register instead of (or together with) survey data, and the role of Central Banks and Eurostat, often have an impact on data accessibility; whether or not this is favourable to access depends on a variety of legal and administrative factors that are strongly country-specific. Another factor that will need to be taken into account in the future is European evolution towards more centralised functioning.

3.1.3 Interpretation of the Law and Practices

The law is often mentioned by NSIs and other governmental agencies as the main obstacle to better microdata dissemination. Recent changes in the European law and the laws of some countries (see above) constitute major improvements, but they are not sufficient as the interpretations of the law and by extension, the practices, often differ across countries and over time. In particular even before the research purpose was introduced in statistical laws, fully anonymised microdata were not considered as falling under the statistical law in some countries, whereas a more restrictive interpretation prevailed elsewhere and no dissemination was allowed. Indeed many NSIs tended to adopt a restrictive interpretation of confidentiality safeguards in order to preserve respondent trust, in the fear that even a single incident may

have a devastating impact especially if it receives media attention. Today, diverse perceptions of confidentiality provisions persist, leading to different requirements for researchers and different degrees of anonymisation. One consequence is that the more restrictive interpretations often tend to prevail at European level for dissemination of Eurostat data sets. In particular, the level of anonymisation of the LFS is higher in Eurostat datasets than in several national samples and is generally considered by researchers as inadequate for many research projects. Nevertheless, many NSIs have been recently paying increasing attention to researchers' needs, especially in countries where confidence has gradually been built and transparent procedures for access have been designed. Best practices in the relationships between Data Archives and NSIs at national level set the example for building a new partnership between the cessda-ERIC, Eurostat and NSIs in Europe.

In sum, differences in the interpretation of the law and in actual practices have often resulted in discrepancies in access across countries. While most NSIs tend to adopt a conservative interpretation, the process of building confidence that is going on in some countries has opened the way to improvements in access for researchers.

3.2 Current access arrangements

This sub-section reviews general information on current access arrangements to governmental statistical microdata across the European research area. Table 1 below summarizes this information. No distinction is made, at this stage, among the different types of data producers –whether it is the NSI or other governmental agencies. Similarly, the table does not take into account the type of distributors, i.e. whether the data producer or some other organisation (notably a CESSDA Data Archive) is in charge of dissemination. The question of whether a Data Archive can constitute an appropriate or even a safe channel for access is not dealt with at this stage, but will be taken up in the next sub-section.

Each cell indicates the rough number of countries adopting a given access arrangement for a given type of data –for instance, cell 1.1 shows approximately how many countries have Public Use Files (PUFs) for population census data. Lighter shades of blue correspond to a smaller number of countries, while darker shades of blue correspond to a higher number of countries.

The first impression is that modes of access and the types of files to which access is granted vary widely across European countries. Looking at the situation country-by-country might reveal even greater diversity, to the extent that the same type of data may be unequally accessible if different public-sector agencies are involved in its production and/or dissemination, especially when central coordination of the national statistical system is weak.

Despite this diversity, some common trends can be detected that can be of interest for a future European Research Infrastructure. Let us comment the table with respect to the different rows (types of data), then with respect to columns (modes of access).

Table 1: Current access arrangements to government statistics throughout Europe

	1. Public Use Files	2. Scientific Use Files	3. Extracts (subsets)	4. Public tabulations	5. Special (bespoke) tabulations	6. Secure remote access / execution*	7. On-site safe centres
1. Population Census, Register, or Microcensus							
2. Main Household Surveys							
3. Some Data from Administrative Registries							
4. Some Business Data							
5. Some other economic and financial data							
6. Other							

3.2.1 Types of data

Rows 1 and 2 indicate that there has been generalised progress in access to anonymised household survey data and to population census and microcensus data. Remaining gaps in access largely reflect differences in the way data collection and dissemination are organised nationally: regarding row 1 (census), the Nordic countries rely largely on register data, while most other European countries conduct questionnaire-based censuses, and Germany has had only microcensus for long[6]. There may be further improvements in the near future: for instance, Scientific Use Files (SUFs) for census are not yet available in Italy, but forthcoming. Regarding row 2 (household surveys), differences between data producers are likely to be a source of heterogeneity even if only the most important surveys are considered (Labour Force Survey, Household Income and Expenditure Survey, Household Panel Surveys, Youth Cohort and Education Longitudinal Studies). Future improvements can be expected: in particular, PUFs for Swiss surveys are not yet entirely available but in the process of being prepared.

Access to administrative data (row 3) is generally limited, with heterogeneities that partly reflect disparities between the structures of national statistical systems: first, some countries

[6] The last German census was conducted in 1987, i.e. before reunification; the next census is planned to take place in 2011 and will be based primarily on register data. The Microcensus is a representative 1% population sample of the German population and has been collected since 1957 in the former West Germany and West Berlin, and extended to the former GDR in 1991. The European Labour Force survey (EU-LFS) is a 0.45% sub-sample of the Microcensus.

have more experience with register data than others; second, the dissemination of administrative data is sometimes entirely in the hands of the producer (central registry or other governmental agency), while in other cases, it is mediated by the NSI which centralises management of this kind of data. For example SORS, the Slovenian NSI, is authorised to retain any government-produced data that it may need for reporting purposes, so that it can act as a common point of access for data produced by a variety of government agencies.

Access to business data (row 4) is usually limited, and raises concerns especially in small countries (e.g. Austria) where data are difficult to anonymise. SUFs, Bespoke tabulations, and on-site safe centres are the means through which at least some countries make some of their business data accessible to researchers. This category also includes data on agricultural firms and the European Community Innovation Survey, where they are made available.

The category of "other economic and financial data" (row 5) includes banking and financial data, tax data, and external trade data, to the extent that they are made available to researchers.

On the whole, access to anonymised data sets has generally improved, especially as concerns major household surveys and, to a lesser extent, census. Access to other types of data, particularly administrative, business and banking data, is less developed and displays a greater degree of heterogeneity across Europe.

3.2.2 Modes of access

Availability of Public Use Files (PUFs) and Scientific Use Files (SUFs) is presented in columns 1 and 2. The distinction between the two modes of access is not always clear and may vary from one country to another, not least because diverse perceptions of confidentiality coexist. In general, it can be said that SUFs are meant for use by accredited researchers only, while PUFs are destined to a larger category of users, also including students (with products to which some countries refer to as "CAMPUS" files, e.g. in the case of Germany). In addition, both PUFs and SUFs are files of anonymised data that exclude direct re-identification of individual respondents, though the degree of detail is different. In particular, some countries adopt the notion of *de facto* anonymisation to distinguish SUFs with respect to PUFs, meaning that indirect re-identification is not entirely impossible but would take such a disproportionately large investment of time, labour, and money that it is most unlikely to occur. Finally, another difference is that data users are often required to sign a contract or licence agreement to use SUFs, while licences are rarely required for PUFs. Yet levels of detail and the interpretation of factual anonymity may still vary from country to country, and may change over time in response to varying expectations on data protection by the authorities and the public. The terminology is also fluctuating; for instance, the category of PUFs covers what is known in Ireland as Anonymised Microdata Files (AMFs), while the category of SUFs covers Research Microdata Files (RMFs).

Column 3 refers to extractions made on request from researchers. This solution is often employed to limit disclosure risks in the case of sensitive data; for example, the Dutch and French NSIs use it for health data such as causes of death.

Columns 4 and 5 indicate that tabulations, both standardised and tailored to the specific needs of individual users, are also very popular though they only provide descriptive information. Their success has been recently enhanced by the development of web-based tools to prepare and customise tables (see section 3.4). Tabulations are also used as one way to provide data without enabling access to confidential information. For instance, since 1999 French census subsets no longer include details on country of birth and citizenship, but this information can

be made available through tabulations (mostly bespoke tabulations tailored to the specific needs of researchers, especially when geographical details are also requested).

Finally, secure solutions for access to confidential data, whether on-site or remote (columns 6 and 7), are increasingly popular, and can be expected to spread significantly in the next few years. On-site safe centres allow researchers to access very detailed microdata on the premises of NSIs or other producers of governmental statistical data, under secure conditions. Usually in a dedicated room, they are allowed to use a special IT environment with no downloading or e-mail facilities; intermediate printing is often allowed but only final outputs can be taken out by researchers once results have been checked for confidentiality by NSI staff. The NSIs of Finland, France, Germany, Hungary, Ireland, Italy, the Netherlands, Slovenia, Spain, and the UK have this mode of access. Austria is currently conducting a feasibility study on a Safe Centre project; instead, Denmark used to have an onsite safe centre, but has now discontinued it to rely exclusively on its secure remote connection facility. Other, non-NSI producers of governmental statistics sometimes have safe data centres as well: for instance in Germany, BA-IAB (the research and statistics service of the Federal Employment Agency), and in Norway, the Norwegian Directorate of Immigration (UDI).

Secure remote access solutions consist in the use of a special IT environment with enhanced firewall protection, no availability of printing, downloading or e-mail facilities, control of inputs, and output confidentiality checks; the advantage with respect to on-site arrangements is that researchers do not have to be physically at the NSI's premises but can use the data from their own workstations. Broadly speaking, there are two types of secure remote facilities, one that can be called "remote access" and allows researchers to see the data they are working with, and "remote execution" which does not allow researchers to actually see the data: they can only submit job requests by email, wait for the analyses to be performed by NSI staff, and receive the results by email after a confidentiality check. Remote access facilities also differ in the actions they allow and the frequency of output checks (all outputs or a random selection), despite the fact that logs are always conserved for possible future control. Secure remote execution systems are available at the NSIs of Austria, Denmark, Finland, Germany, the Netherlands, Slovenia, and Sweden; secure remote access is offered in France (as a pilot at the moment), Ireland, the Netherlands, and is planned to be operating soon in the UK (also as a pilot). Statistics Netherlands is the only one to offer both types. Non-NSI producers of governmental statistics sometimes have such facilities as well: it is the case, for instance, of the Bank of Italy (BIRD system), BA-IAB in Germany, and LIS (Luxembourg Income Study) who all offer remote execution systems[7].

In general, safe centres were created by the NSIs who first moved in the direction of offering greater access to confidential data, while secure remote solutions have been chosen by most late-comers. This is not always true, however, as shown by the case of Austria that first implemented a secure remote execution system and is now considering an on-site safe centre, and by the case of Hungary whose "Research Room" is of very recent creation (September 2008). The case of Denmark indicates that remote connection facilities may eventually supersede on-site safe centres; yet the case of the Netherlands also hints that the two types of solutions may well coexist as they serve different purposes (remote access for more experienced users, on-site facilities for those who need support or advice from staff). Germany's solution also combines on-site facilities and remote access connection, a system that was established to overcome problems of access due to restrictive legislation. There are

[7] More detailed technical, organisational, and legal information on secure remote access systems is available in the WP10 tendered report on "Secure Remote Access system for an upgraded CESSDA RI", prepared by Metadata technologies on behalf of CESSDA.

some signs that very sensitive data may be only available through on-site safe centres for a long time in future. With respect to secure remote connection facilities, on-site safe centres have both advantages and disadvantages for researchers: indeed despite the cost of travelling and lodging at the safe centre's location, they offer help from staff which is sometimes a crucial resource for research.

On the whole it can be said that, although PUFs, SUFs and tables still constitute important and widespread forms of access to data, the most dynamic area is that of modes of access to confidential data supported by secure IT systems, both onsite and remote. Further developments and extension of use of these systems is to be expected in the years to come.

3.3 Costs

This sub-section reviews costs which constitute one of the main existing barriers to access. Costs are particularly interesting in view of a future ERIC because the pressing need to share costs might favour enhanced collaboration between NSIs and the research community with the mediation of CESSDA Data Archives.

The importance of costs arises from the fact that, leaving aside privacy protection issues, most remaining barriers to access could be removed with greater funding. For instance, though technical barriers are losing part of their relevance due to the rapid development of IT tools, they are still meaningful for small countries which cannot afford up-to-date IT infrastructure. Further, investments in documentation are needed in many countries because existing metadata are not always sufficient for secondary analysis and because availability of harmonised documentation for comparative research is unsystematic.

Yet costs remain an unsolved issue for many data producers, even if they charge fees to users. Fees are often insufficient to recover costs, and an increasing demand by researchers often imposes a strain on NSI staff. This is because only some types of unit costs are decreasing in the number of users (e.g. the cost of documenting data), while others are increasing (e.g. the cost of performing confidentiality checks on researchers' output at safe centres).

In general, the current debate over costs is dominated by two lines of argument. On one hand, it is argued that data produced with public funds should be provided free of charge not only to researchers but to all interested users. A tendency to reduce user fees is present in numerous countries and can be expected also for Eurostat datasets. On the other hand, the rising importance of specific services, especially those related to access to confidential data, tend to generate new and significant costs, so that one option currently under consideration is sharing these costs between government data providers and research institutions (research ministry/council, data archives, research institutes/universities, etc.).

At present, most data producers charge fees to users. In most cases, the data itself is free of charge and fees are meant to cover the extra expenses to make it available, in particular to extract subsets on demand by researchers, to prepare bespoke tabulations, to link data, and in general, to satisfy individual needs for which standard products are insufficient. Charges are also due for data delivery and in some cases, for partial coverage of documentation costs. Finally in the case of sensitive data (on-site safe centres and secure remote connections) charges are often due for output checks. Some examples of end user fees are listed in Table 2; however, not all countries have a pre-specified charging policy, and many determine end user fees on a case-by-case basis. It should also be added that end users may face indirect costs together with fees: in particular secure data centres on the premises of NSIs are expensive for users even when no fees are charged, because of the travel and subsistence costs that must be incurred to use them when no provisions exist to cover these costs.

Table 2: examples of end user fees

Country	Fees
Czech Republic	– Bespoke tabulations: (about) 16.50 euros; – Household Budget survey (1 year): 825.00 euros; – Microcensus 2002: 618.00 euros; – Microdata from innovation survey: 62.00 euros per reference period. – Other surveys: individual calculations apply.
Finland	"Astika" (times series database with statistical data on the Finnish economy): – 2020.00 euros + VAT 22% for one user, 6055.00 euros + VAT 22% for one organization (3 or more users); – the price per month for one user is 404.00 euros + VAT 22%.
Germany	– PUFs: fee which varies according to the type of file. – CAMPUS-files (a type of PUF): no cost. – SUFs: 95.00 euros; – Secure remote connection: 95.00 euros; – Use of on-site safe centres: 95.00 euros.
Italy	– Census: 150.00 euros + VAT 20% per file; – NSI (ISTAT) surveys: 90.00 euros + VAT 20% per file. – Use of on-site safe centre: free of charge.
Netherlands	For remote access and use of on-site facilities at Statistics Netherlands: – Fixed cost per user: 170.00 euros configuration of user, plus 450.00 configuration of workstation for remote access); – Fixed administrative cost per project (150.00 euros); – Variable cost depending on time needed: 46.00 euros per half-day in on-site safe centre; 480.00 euros per month for remote access; 170.00 for output check. Remote execution: – 92.00 euros or 129.00 per job, depending on the software used. Additional charges are due for data preparation and documentation.
Norway	Extractions and linked datasets: 50.00 - 100.00 + VAT per hour.

Where fees exist, they are dependent upon the legal framework at least in part, in cases in which the legislation states that statistical agencies must fully or partly recover costs from users. The internal regulations of NSIs and agreements with the Ministry of Research and/or Education also play a role in some cases, especially when the Ministry or some other national research agency agrees to cover part of the costs on behalf of the whole scientific community. Table 3 summarizes how costs are distributed among the different stakeholders in Europe. As in Table 1 above, lighter shades of blue correspond to a low number of countries, while darker shades of blue correspond to a higher number of countries, adopting a given solution. It appears that although most of the burden falls on data producers and end users, national agencies sometimes cover at least part of these costs for the whole research community. For instance, the Danish Ministry of Research allocates 800,000 euros a year to Statistics Denmark to reduce the costs of its online remote execution system. Similarly, the German Ministry of Education and Research funded the creation and the first years of operation of the Research Data Centres of the Federal Statistical Office and the Statistical Offices of the States, of the Employment Agency, and (partially) of the Pension Fund and of the Federal

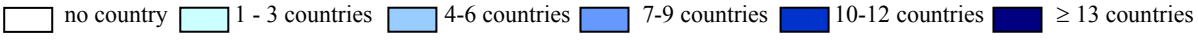
Agency for Vocational Education. In the UK, the Economic and Social Research Council is currently funding a secure remote access system that is going to be implemented by UKDA, the CESSDA member, as a pilot project. In France, the CESSDA Data Archive (Réseau Quetelet) covers costs of access for researchers to specific datasets (SUF or tabulations) produced by INSEE, the National Statistical Institute; an agreement to share costs for remote access is currently being negotiated. In Germany, CESSDA member GESIS (GML unit) also shares some costs with government data producers, and participates in a number of co-funded projects particularly for anonymisation and documentation. ADPSS, the CESSDA member for Italy, covers costs of access to data from ISTAT (the Italian NSI) for its users; however, ADPSS is only allowed to distribute ISTAT data to members of the Department of Sociology and Social Research at the University of Milan-Bicocca of which it is part, so that despite its being formally a CESSDA member, it is in fact more similar to the case of a higher education/research institution covering costs for its affiliated members only. Other examples may be provided but a general warning is necessary, because cost arrangements may vary within the same country, if e.g. the Ministry covers costs for researchers' access to NSI data but not to administrative data produced by, say, the unemployment or pension agency.

At European level, Eurostat charges fees for release of microdata to researchers. This is partly due to European regulations, and partly to the fact that these data are only available at a fee in some countries. Though charges have recently decreased, they still constitute a major barrier for researchers: they are to be paid for each project and require the signature of a contract with the research institute or university of affiliation of the researcher. Till now, some contracts have been signed for European research networks (see for instance EQUALSOC) but no national arrangement has been possible to cover access for all researchers.

On the whole, this suggests that CESSDA may play a more important role in this area.

Table 3. Who covers costs of access to government data in Europe

	Provided by the producer, free of charge for users	Fees paid by end users	Charges are paid by a higher education or research institution, for all affiliated members	Charges are paid by Data Archive, Ministry, or National Research Council, for the whole scientific community
Public use files				
Scientific use files				
Extracts (subsets)				
Public tabulations				
Specific (bespoke) tabulations				
Secure remote access / execution				
On-site access through safe centres				



In sum, it can be said that costs constitute one of the main remaining barriers to access, together with legislation and practices. They are a major source of concern both for the supply side (NSIs and other data producers) and for the demand side (researchers). CESSDA organisations already contribute to alleviating the financial burden in some countries, at least for NSI data; they may consider doing so more systematically all over Europe.

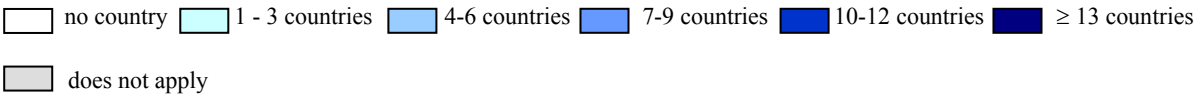
3.4 Levels of Access, and Providers of Access

This sub-section develops the question of who are the providers of access, together with the related question of levels of access. Here, the question of the role of CESSDA Data Archives as providers of access to governmental statistics is addressed explicitly.

The following table summarizes information collected to address this issue, by modes of access. It distinguishes government data producers according to the extent to which they use the world wide web, email, and other forms of online connection (column 1) with respect to more traditional means such as regular mail or onsite access on their premises (column 2). The table also indicates the degree to which government data can be accessed from CESSDA Data Archives (column 3) or from comparable data organisations which are not currently members of CESSDA. Again, this is a general overview that may mask within-country variation: indeed, agreements between CESSDA data archives on the one hand and NSIs and/or other agencies on the other hand, may widely differ.

Table 4: Level and Providers of Access

	Access granted directly by producer through the Internet	Access granted directly by producer in other form (onsite, post)	Access from a CESSDA data archive	Access from a non-CESSDA distributor
Public Use Files				
Scientific Use Files				
Extracts (subsets)				
Public tabulations				
Bespoke tabulations				
Secure remote access / remote execution				
Onsite safe centre				



3.4.1 The effects of the Internet on NSIs' dissemination practices

The table shows the increasingly important role of web-based tools for dissemination, especially for the preparation of tables. For instance in the Netherlands, public tabulations are provided through the web-based application "Statline" (http://statline.cbs.nl/statweb/) developed by CBS, the Central Bureau of Statistics. Similarly, Statistics Norway has developed "StatBank Norway" which allows users to select scope and content of each table, and then to export results in several data file formats (http://statbank.ssb.no/statistikkbanken/default_fr.asp?PLanguage=1). A similar tool exists in Denmark (http://www.statbank.dk/statbank5a/default.asp?w=1024). Other NSIs are currently investing to improve their web-based access to tables. Statistics Austria expects to replace its old online tabulation system ISIS with the new database system "SuperSTAR" (http://www.statistik.at/web_en/publications_services/superstar_database/index.html), which is currently being tested. In Greece, the National Statistical Service is currently updating its website (http://www.statistics.gr/StatMenu_eng.asp), aiming to use it to harmonise its data dissemination policies with European recommendations on official statistics.

The Internet has also brought about rapid changes in the distribution of PUFs. Once difficult to obtain, PUFs can often be downloaded directly from NSIs' websites now, though this solution is still less widely adopted than more traditional ones. In Germany for instance, only CAMPUS files are freely downloadable, while other PUFs are sent to users via postal services on CD support. The latter solution is preferred by several other countries such as the Netherlands and Norway. Other countries adopt a mix of on-line and off-line modes of distribution; for instance, Italy's ISTAT sends bespoke tabulations via email and does the same for other data files, recurring to CDs only if their size does not allow email transfer; ISTAT is also planning to make SUFs downloadable from its own web site.

To summarize, the Internet has brought about a generalized improvement in availability and accessibility of anonymised data released by the NSIs themselves, particularly tables and, to a lesser extent, PUFs.

3.4.2 CESSDA Data Archives as distributors of government data

The table also shows that CESSDA Data Archives play a major role in the dissemination of governmental data to researchers. Eleven members currently hold at least some government data, namely:

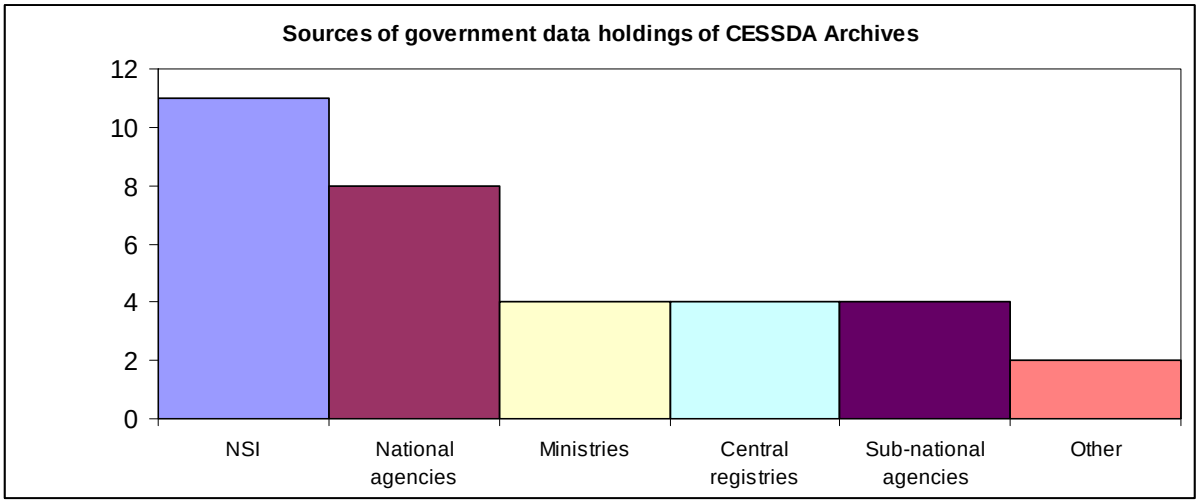
- ADP (Slovenia),
- ADPSS (Italy),
- CEPS/INSTEAD (Luxembourg),
- CNRS-RQ (France),
- DANS (the Netherlands),
- EKKE-GSDB (Greece),
- GESIS (Germany),

- ISSDA (Ireland),
- NSD (Norway),
- TARKI (Hungary),
- UKDA (UK).

Eight of them (ADP, CEPS/INSTEAD, EKKE-GSDB, GESIS, ISSDA, NSD, RQ and UKDA) have a formal agreement with their NSI, and some of them also with other producers of governmental data; FORS (Switzerland) has recently reached an agreement with the Federal Statistical Office. However, the quantity and quality of governmental data holdings of CESSDA organisations vary widely: some only have aggregate data, while others do have microdata but are not allowed to distribute them as such, and can use them only for preparing tables, SUFs/PUFs, and documentation. In fact only a restricted subgroup of CESSDA organisations has experience with microdata distribution. This is reflected in the fact that many CESSDA Data Archives wish to include NSI data in their collections, or alternatively to increase their holdings of NSI data. The survey that was distributed by the PPP team to the full CESSDA membership in the summer 2008 included a question on the types of data that they would like to have if they do not already form part of their collections, and NSI data appeared to be at the top of the list, with six organisations declaring they would like to include them in their collections or to increase the amount they have. These organisations are: FSD (Finland), FORS, GESIS, RODA (Romania), SND (Sweden), and TARKI.

The following chart distinguishes the sources of governmental data holdings of the 11 CESSDA organisations that do have them. NSIs are the most important data providers: all those that have government data include at least some datasets provided by their NSI. Other public sector organisations are present as well, though with greater heterogeneity across countries. They include public-sector agencies at national level (e.g. tax administration, health and social security agencies, etc.), ministries or government departments, public-sector bodies at sub-national level (region or town), and other types of governmental institutions. The picture must be completed by taking into account differences in the structure of national statistical systems, e.g. the fact that Central Registries may not exist in some countries, and that the NSI sometimes centralises the management of administrative and other types of data. For example in Norway, NSD has microdata from Statistics Norway, Ministries, Central Registries, and parishes; in the UK, the holdings of UKDA include, together with data from the Office for National Statistics (ONS), also data produced by the Home Office, the Department for Work and Pensions, the Department of Business, Enterprise & Regulatory Reform, the Department of Environment, Food and Rural Affairs (DEFRA); sub-national data are supplied to UKDA by ONS.

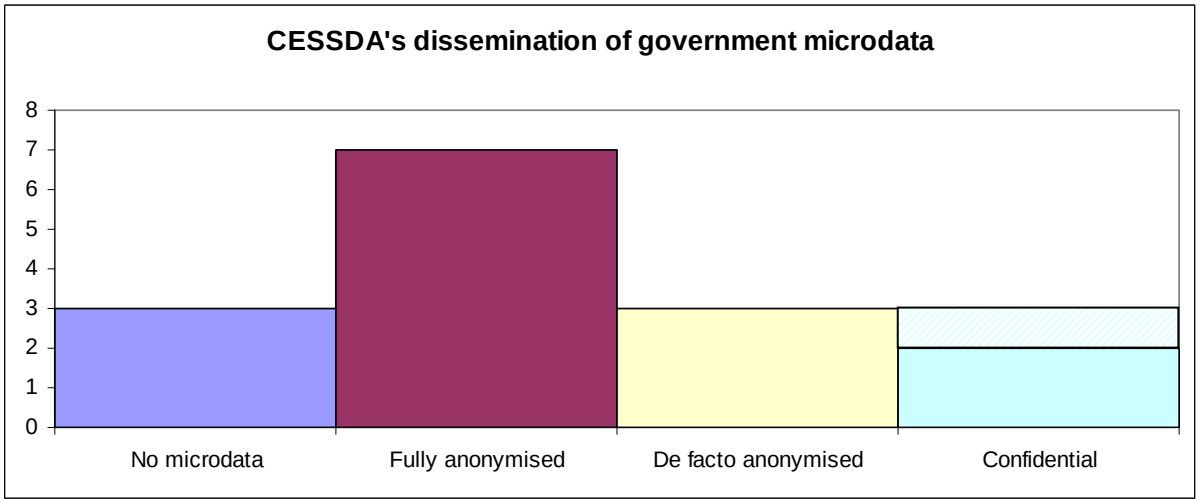
Figure 1



The following chart provides further information on levels of detail available for microdata distributed by the 11 CESSDA members that hold some type of government data:

- Three of them only distribute NSI aggregate data and/or tables but no microdata, namely ADPSS (Italy), EKKE-GSDB (Greece), and TARKI (Hungary).
- Fully anonymised microdata are distributed by ADP (Slovenia), CEPS/INSTEAD (Luxembourg), CNRS-RQ (France), DANS (Netherlands), ISSDA (Ireland), NSD (Norway) and UKDA (UK).
- CNRS-RQ, NSD and UKDA are also involved in the dissemination of factually anonymised data. For instance RQ distributes SUFs from INSEE, the National Statistical Institute, and offers support for access to more detailed data, in particular by programming computer code to prepare bespoke tabulations for users from INSEE data; it also collaborates to an ongoing pilot project aimed at setting up a secure remote access facility for access of researchers to public-sector confidential data.
- Finally, NSD and UKDA are also actively involved in the distribution of confidential data. For example, NSD manages access to the KIRUT database, built by merging register data concerning the welfare system and the labour market. The GML unit of GESIS is included in the "confidential" area as a particular case (a shade instead of a full colour) because it is not normally allowed to distribute SUFs, although under some very specific conditions, it may offer access to detailed datasets on its own premises in Mannheim. This procedure is adopted in cases in which researchers need datasets for which SUF versions are unavailable, and in the case of foreign researchers who are not allowed to receive SUFs (see section 3.5.4).

Figure 2



To conclude, a short note on the non-CESSDA data organisations that play a role in the dissemination of government data is necessary. These organisations are not numerous but may be important in some countries. For instance in Germany, IZA (Institut für die Zukunft der Arbeit, Institute for the Study of Labour) manages the system ISDC (International Data Service Centre) which offers onsite access and secure remote execution. In Spain, some data produced by INE (the Spanish National Statistical Institute) are made available in the form of PUFs by a non-CESSDA Data Archive, CEACS – Instituto Juan March. As mentioned in the Methodology section some areas, particularly health and geo-tagged data, are part in some cases of a parallel access system not integrated with CESSDA, although researchers increasingly demand to combine data provided by both systems. This is a matter that calls for further reflection in view of the future planning of the ERIC.

Finally, it must be reminded that no CESSDA member is currently entitled to disseminate Eurostat microdata. Contracts with Eurostat are signed individually with research institutions and universities for each specific project, an arrangement that is widely perceived as a major barrier throughout Europe.

In sum, CESSDA members play a major role in the dissemination of government data. They are primarily active in the dissemination of aggregate data, tables, and anonymised microdata, but some of them also offer access to sensitive data. There are disparities across countries, though, with some members who have a major role in government data dissemination and others that have hardly any relationship with statistical agencies. The fact that many of them wish to include government data in their collections, or to enlarge their current holdings of government data, is an opportunity to fill this gap and to enhance the role of the whole CESSDA as a provider of government data. If the future ERI considers this opportunity, however, it will need to take into account existing forms of distribution through non-CESSDA organisations.

3.5 Researchers' accreditation

This sub-section is about researchers' accreditation, with special emphasis on the need for some data to cross borders in order to make comparative research possible at least within Europe, and on the need to provide sufficient data training to future researchers.

The question of accreditation is all the more important as the fact that research gains increasing recognition in the legal framework does not always answer questions such as: who is a researcher, and what is a research. These two questions must be answered simultaneously because a professional researcher may use data for other purposes (e.g. consultancy for business), while non-professional researchers (including students) sometimes engage in research activities. Broadly speaking, however, the accreditation process is not limited to defining criteria to identify research and researchers, but consists of a multiplicity of tasks also including implementation of the criteria and management of applications. A variety of actors can be involved, ranging from the legislator to the NSI, statistical authorities, CESSDA data archives, independent committees, etc.

3.5.1 Who is a researcher, and what is research

The question of who is a researcher often receives an institutional answer: researchers are those who are affiliated to a recognised university or other higher education/research institute. The general principle is that researchers are submitted to regular internal evaluation procedures in these institutions. In some countries (e.g. the Netherlands, Norway), the full list of acknowledged institutions is stated in the Statistical Law, with a provision allowing non-listed institutions to obtain recognition under some conditions. At European level, Eurostat also considers a list of institutions provided by Member States. In other countries, the status of researcher is assessed on a case-by-case basis; in particular, a research project is often required, in which the data user clearly indicates not only the purpose of the research, its expected output, its duration, methodology, and the full team of collaborators, but also states the variables needed and how they will be used. In other cases, additional indicators are also used to qualify an applicant as a researcher, such as a list of publications, paper presentations, or past research experience; the new "Approved Researcher" procedure in the UK also invites data users to indicate their previous experiences with detailed microdata as a support to their application. Often, though not always, requests for access to confidential data are processed individually, whereas access to anonymised data (PUFs, SUFs, and tabulations) relies on simpler procedures; the amount of information required typically differs in the two cases. This diversity of solutions results in strong heterogeneity across Europe, reinforced by the fact that information on accreditation criteria and procedures is not always easy to locate on the web. In this context, the new ERIC may consider offering solutions to centralise provision of information about criteria and procedures at national and European levels.

Evaluations of what is a research have traditionally been based on a set of parameters: source of research funding, purpose of project (scientific publication and sometimes also PhD dissertation or Bachelor/Master's thesis), absence of commercial use. An issue that a few statistical systems are currently addressing is that public policy evaluation activities blur the frontiers between research and governmental statistics.

Once accreditation has been granted, the researcher (and in some cases, a representative of his or her institution) is often required to sign a licence agreement for the use of data. The rationale of these rules is that research institutions are often held responsible for data use together with individual researchers, and can be sued in case of breach. In Norway for example, permission/licence to collect, access, and/or process confidential information is

given upon condition that the institution takes formal responsibility for the data; this rule applies to all research projects involving personal information.

Conditions typically include the use of the data for a single project only, a requirement to report all publications, prohibition to cede user rights to third parties; sometimes, researchers also have to destroy data after use, or submit papers for confidentiality checks before publication. Though with a common core, such conditions vary across countries and within each country, they sometimes vary according to the type, the producer, and the level of detail of the data. Many models exist, ranging from a short list in a paragraph to several pages with details on all conditions. Again, differences between countries are accompanied by differences within the same country.

In sum, recognised research institutions are often mentioned in the law but there are major cross-country differences in the criteria to define a researcher and a research activity as well as in requirements for accreditation.

3.5.2 Implementation of criteria and management of applications

Within this framework, procedures for the practical implementation of criteria and management of applications often involve a committee including representatives of researchers. Such committees sometimes have only advisory functions, sometimes fully participate in the decision-making process. Differences typically exist for anonymised and confidential data. For instance in the UK, accreditation decisions are the responsibility of the Statistics Authority, which includes representatives of researchers. In Portugal, a scientific committee of the Ministry of Research advises the NSI (INE) about accreditation issues. Some CESSDA organisations take responsibility for accreditation decisions for anonymised data at least in part, once general conditions have been agreed with the governmental data provider. A case in point is France, where general principles for PUFs and SUFs are agreed at national level between the NSI, the other government statistical services and the Ministry of Higher Education and Research (*Comité de concertation pour les données en sciences humaines et socials*, CCDSHS) on the basis of recommendations from a Scientific Council, and accreditation is devoted to the CESSDA member Réseau Quetelet. The latter can seek advice from the Scientific Council in problematic cases. Accreditation for confidential data involves a different committee, the *Comité du secret statistique*, which acts in the frame of the statistical law. Since 2009, researchers are involved in the process through direct representation of the Scientific Council of the CCDSHS in the *Comité du secret*[8].

Another example is the UK, where the CESSDA Data Archive UKDA acts as an intermediary for Approved Researcher status application in the case of confidential data, while the decision is made by the Statistics Authority. In the case of confidential data, the Norwegian Data Protection Official at NSD (the CESSDA member), on behalf of the Universities, colleges and research institutions, has the authority to approve research projects that meet the requirements of the Personal Data Act and the Health Register Act. The process of reviewing researchers' applications with regard to the relevant data protection laws is additional and not substituting the accreditation process by the data producer and data providers; the latter maintain an independent responsibility to establish whether this is research and whether the researcher is affiliated to an accredited research institution. Finally in Switzerland, the Federal Statistical Office takes full responsibility for accreditation decisions, but in the new framework of collaboration with the CESSDA member FORS, the latter offers information in

[8] Accreditation needs to be confirmed by the CNIL, the authority in charge of enforcing the privacy protection law, and by the National Archives.

support of the Office's decision, based on a database of research in Switzerland that FORS maintains.

In the case of Eurostat, only NSIs are involved, with no formal participation of representatives of the research community. Eurostat's accreditation process is based on a set of general, pre-defined principles (list of trusted institutions provided by each member state).

Implementation of accreditation criteria and management of accreditation applications often involve representatives of the research community. In some cases, CESSDA members take full or partial responsibility for accreditation procedures, while in other cases they perform an advisory function.

3.5.3 Accreditation of students

Accreditation issues raise additional problems relative to the data needs of students. Future researchers should receive adequate training with microdata, but students cannot offer the same guarantees as confirmed researchers. Not only do they lack the record of publications and research experience that researchers are sometimes asked to exhibit in order to obtain access, but they are not always employed by a research institution which could guarantee for them. Solutions vary: while some countries offer students the same access opportunities as confirmed researchers (e.g. Norway), others distinguish between undergraduates and graduates and within the latter group, between Master and PhD students. In this perspective, students' access is often restricted to PUFs and tables, while PhD candidates (and sometimes, also master's students) can obtain more detailed data. Other countries impose additional requirements, e.g. the thesis supervisor must take responsibility for the student's data use.

For instance, German students are generally expected to use PUFs, more precisely CAMPUS-files; however, they may be allowed to receive SUFs if their work (PhD dissertation or master thesis) is part of a wider research project conducted by their Institute or Department, and requests for access are made by the University or Higher Education institution they belong to. In Hungary, students can only carry out onsite research. Italian students are only allowed to access ISTAT's PUFs, while they are considered as no different from confirmed researchers by the Bank of Italy. In the Netherlands, only employees of the research and higher education institutions acknowledged in the Statistical Act can access data; students are allowed to access data if they also satisfy this condition, for instance through a temporary employment contract with their institution.

Across Europe, accreditation conditions vary more widely for students than for senior researchers.

3.5.4 Accreditation of foreign researchers

The European Union has adopted a principle of non-discrimination and a uniform legislative framework for data protection, which implies that in principle, data can circulate within the European Research Area. On this basis, the French Social Science Data Committee has authorised foreign researchers' access to French data; the only difference with respect to national researchers is the amount of information demanded in application forms, due to uneven knowledge of foreign higher education and research institutions. In this perspective, CESSDA may play a role by centralising this kind of information and constituting a single point of contact for the different accreditation institutions that may need it. Based on current regulations, a principle of reciprocity may also be applied.

Still, several types of data are in fact not allowed to cross borders, even within Europe. Only PUFs and public tabulations circulate freely; in many other cases, foreign researchers cannot be easily accredited and are able to obtain the data of a country only through a visiting period in a research institution in that country, with obvious repercussions on the costs of comparative research.

For instance in Denmark, remote access is only accessible to foreigners who are visitors at Danish research institutions. In Germany, foreign researchers are not allowed to obtain data in the form of SUFs, but they can use PUFs, on-site safe centres, and secure remote connection facilities. In Hungary, foreign researchers can only carry out onsite research. In the Netherlands, foreign researchers need to have their organisation approved by the Central Commissions for Statistics, a statistical advisory committee. Instead, Italy and Slovenia apply the same rules for foreign and national researchers. Similarly in Norway, foreign researchers are defined as part of NSD's primary user group and are allowed to obtain survey data from Statistics Norway, stored at NSD and disseminated through NSD, at the same conditions as Norwegian researchers. The situation with regard to data from administrative registers varies according to the institutional and legal framework being more or less restrictive. Finally, Sweden entitles EU researchers under the same conditions of access as national researchers, while other foreigners can only access PUFs.

On the whole, it appears not only that there is great heterogeneity across Europe, but also that little distinction is made between foreign researchers from the EU and from non-EU countries. This is surprising in light of the legislative framework which in principle allows intra-European circulation while preventing data to be sent to countries that lack a comparable data protection regulation.

The table below summarizes information on accreditation conditions for students and for foreign researchers, both from the EU and from outside countries.

Table 5: access conditions for students and foreign researchers

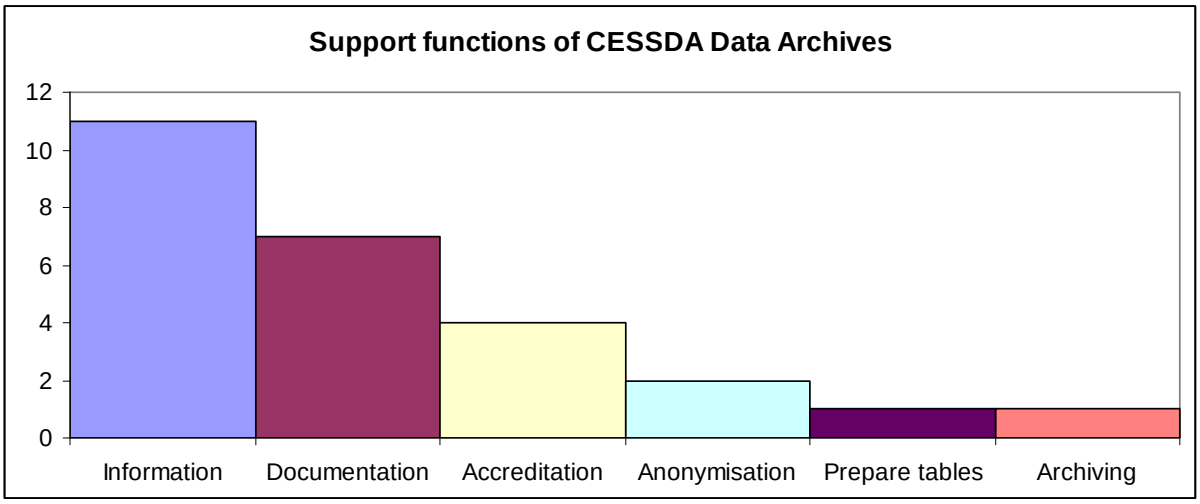
	Same access as national researchers	Access restricted to some types of data files	Additional requirements for access	No access
Students				
Foreign researchers, residents of EU countries (or countries recognized by the EU for data protection)				
Foreign researchers, residents of non-EU countries				

≤ 3 countries 4-6 countries 7-9 countries 10-12 countries ≥ 13 countries

3.6 Functions of CESSDA organisations

Besides *direct* dissemination, some CESSDA organisations offer forms of *indirect* support to users who need to access government data that do not form part of their collections. The above section has already provided an example of such support functions by illustrating CESSDA members' involvement in accreditation decisions and procedures. The following chart shows that CESSDA organisations offer a wider variety of support services to users.

Figure 3



Thirteen out of twenty CESSDA organisations provide some form of support to users who need to access data directly from the NSI (or other governmental producers); four of them are paid for these services. In greater detail:

- Information on access arrangements managed directly by data producers and/or help in data discovery are provided by:
  - ADP (Slovenia),
  - ADPSS (Italy),
  - CNRS-RQ (France),
  - DANS (Netherlands),
  - DDA (Denmark),
  - FORS (Switzerland),
  - GESIS (Germany),
  - ISSDA (Ireland),
  - SDA (Czech Republic),
  - SND (Sweden),
  - UKDA (UK).

  Some of them such as SDA lack a systematic framework for support but do it occasionally, while others offer a more systematic service.
- Documentation services are provided by:
  - ADPSS (Italy),

- CNRS-RQ (France),
- EKKE-GSDB (Greece),
- FSD (Finland),
- GESIS (Germany),
- NSD (Norway),
- UKDA (UK).

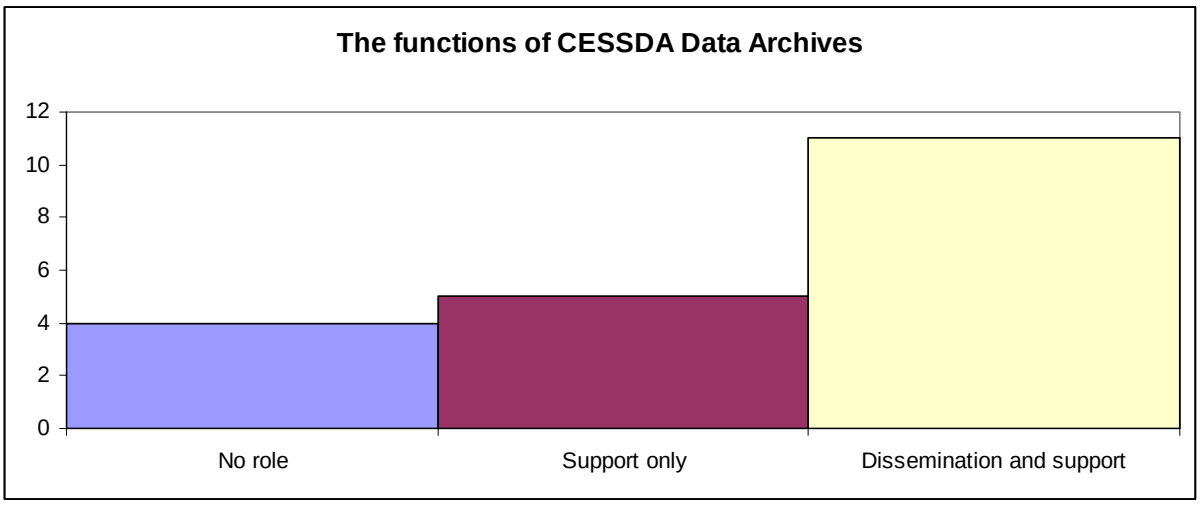
For instance, GESIS's unit GML provides data documentation on the web through its information system MISSY (http://www.gesis.org/Dauerbeobachtung/GML/MISSY/). Expertise in the development of tools for statistical disclosure control and data anonymisation has been developed at GML and at NSD (Norway). In particular GML participated in several projects on anonymisation techniques, often jointly with the German Federal Statistical Office, and received governmental funding for this activity. Its services include preparation of SUFs and of tools for data analysis, personalized consulting services, organisation of workshops and training programs.

- Preparation of computer code to create bespoke tabulations for sensitive data is developed by France's CNRS-RQ staff; the tables are then executed by staff at INSEE, the National Statistical Institute.
- UKDA is an official place of repository for the National Archives of the UK.

This suggests that even in cases in which the legislative framework makes it difficult for a CESSDA organisation to disseminate government data, provision of indirect services for users could be a way this organisation might improve its participation in the system and reinforce its role as a central point of contact for researchers who need to access data.

The following chart concludes by summarizing information on the different functions that the 20 CESSDA member archives have with respect to governmental statistics.

Figure 4



More precisely:

- WISDOM (Austria), ESSDA (Estonia), CIS (Spain) and RODA (Romania) currently have no role in the dissemination of government data.

- DDA (Denmark), FORS (Switzerland), FSD (Finland), SDA (Czech Republic), and SND (Sweden) have support functions only.
- ADP (Slovenia), ADPSS (Italy), CEPS/INSTEAD (Luxembourg), CNRS-RQ (France), DANS (Netherlands), EKKE-GSDB (Greece), GESIS (Germany), ISSDA (Ireland), NSD (Norway), TARKI (Hungary), and UKDA (UK) perform both dissemination and support functions, though with significant differences among them.

Besides direct dissemination, CESSDA Data Archives offer a wide range of indirect support services for users of government data, including information, documentation, anonymisation, archiving, and preparation of bespoke tabulations. The amount and regularity of services provided are very heterogeneous across CESSDA, though: a few archives have strong and long-lasting relationships with statistical agencies, while others have very limited functions and a few of them play no role at all in the dissemination of government statistics.

4. Conclusions and recommendations

The following recommendations and conclusions emerge from synthesis of the previously illustrated findings, and are meant to be considered alongside the other components of this report. They aim to provide CESSDA with guidelines for moving forward.

4.1 Main conclusions

The conclusions listed below are based on the evidence presented in the main body of this report, and set a framework for recommendations on how to envisage further progress for CESSDA as a provider of pathways for access to government data at European level. They cover legal constraints, new developments and challenges, the current situation within CESSDA, the need for enhanced cooperation with Eurostat and NSIs, and the possible role of CESSDA with respect to the question of data circulation across Europe.

4.1.1 A more favourable legislative framework

The legislative framework is often regarded as the main remaining barrier to improved microdata access. However, legal systems are progressively becoming more favourable to data access and cross-border exchange, at least within Europe. While the needs of research were hardly taken into account in the statistical laws and the privacy protection laws that were set up, respectively, after World War II and during the Seventies, they are explicitly recognised in many recent amendments to these laws. It can be expected that the process of European integration will encourage an increasing number of countries to move in this direction. Recent initiatives by OECD and UNECE to prepare Principles and Guidelines on access to data from public funding (OECD 2007) and management of confidentiality (UN 2007), have also contributed to promoting a culture of greater openness. This does not necessarily mean that Eurostat and NSIs will be the main or the only actors of dissemination, as the legislation rarely rules on *how* the data can be distributed in practice, or on *who* is in charge of giving access to them. Yet this tendency may imply stricter requirements in terms of technical protection for access to detailed data (see below).

Despite remaining restrictions, the legislation is increasingly less likely to place a constraint on the reinforcement of CESSDA's role in government data dissemination.

4.1.2 Recent improvements and new challenges in the European data landscape

Access to both anonymised and confidential microdata has recently improved in a number of countries. With the advent of the Internet, availability of tables and PUFs has greatly increased, while new modes of access to sensitive data based on secure IT connections are spreading fast throughout Europe. Yet access arrangements are still heterogeneous across countries, due not only to differences in national legislations but also to differences in the interpretation and perception of the law (especially confidentiality provisions) as well as in practices. Further, new problems are emerging in non-traditional areas such as tax, financial, health-related, and environmental data, increasingly used by researchers but typically considered as highly confidential. These data mostly come from administrative sources but are often demanded in conjunction with NSI data, so that access can only be granted through some combination of different infrastructures. Geo-tagged data raise similar problems. In sum, the set of data of interest for researchers has increased in both quality and quantity, thereby raising new challenges for CESSDA.

A generalized improvement in data availability and accessibility is accompanied by uneven conditions of access across countries as well as by new challenges brought about by changes in researchers' demand and by increased production of high-quality administrative data. These new areas and developments in access conditions should be monitored closely by CESSDA if it is to provide pathways for access to all *types of data.*

4.1.3 The current position of CESSDA organisations

CESSDA organisations are important actors in the dissemination of government data, though a remarkable heterogeneity emerges from the analysis. The situation can be summarized as follows:

- CNRS-RQ (France), NSD (Norway), and UKDA have long-term relationships with NSIs and other producers of governmental statistics, reinforced by formal agreements which enable them to participate in the dissemination of varying types of microdata, ranging from fully anonymised to more detailed and even sensitive.
- ADP (Slovenia), CEPS/INSTEAD (Luxembourg) and ISSDA (Ireland) also have good relationships with governmental statistics but are only allowed to distribute fully anonymised data.
- Other CESSDA members that also distribute anonymised data consider this activity to be somewhat secondary with respect to their core business, notably DANS (the Netherlands) and GESIS (Germany); the latter is a particular case because government microdata only form part of the activities of its sub-unit GML, which was integrated into GESIS only recently.
- ADPSS (Italy), EKKE-GSDB (Greece) and TARKI (Hungary) have some relationship with governmental data producers and even formal agreements (in the case of ADPSS) but are not allowed to distribute microdata. In the case of ADPSS, a major obstacle to further expansion in this area is the nature of the organisation which serves primarily the needs of the University Department of which it is part rather than the national research community; in the case of EKKE-GSDB, the current difficulties reflect the

lack of a national framework for access, but recent developments and the NSI's current efforts to meet EU recommendations for data accessibility are likely to lead to improvements in the near future.

- FORS (Switzerland) is a particular case as it is currently in the process of implementing a new agreement with the NSI that may give it a prominent role in government microdata distribution.
- Finally, WISDOM (Austria), CIS (Spain), and RODA (Romania) do no include any government datasets in their collections; others such as DDA (Denmark), FSD (Finland), SDA (Czech Republic), and SND (Sweden) do not hold government data either, but offer some limited support to users.

The reasons for this heterogeneity are varied. Data Archives are of recent creation in some new accession countries, and need time to enlarge their collections. Some NSIs in these countries adopt a restrictive interpretation of confidentiality provisions, while others are more open, allowing for potential improvements. Other CESSDA Data Archives have traditionally focused on data produced by academics (particularly socio-political surveys) and have never invested in their relationships with NSIs. Either way, CESSDA members with little or no involvement in the distribution of government data have limited knowledge about existing access arrangements in their home countries. While some of them would be glad to include government data in their collections (e.g. FSD, RODA, SND), others seem more reluctant to start a process of negotiations with public-sector organisations, and sometimes overestimate the difficulties.

In cases in which CESSDA members take little or no part in the dissemination of government data, NSIs meet researchers' demands themselves, sometimes with financial help from the Ministry of Research (e.g. Denmark, the Netherlands); the arrangements they have set up are rather research-friendly in some countries (e.g. Denmark, the Netherlands, Sweden), but less so in other countries (e.g. Greece, Romania). In some cases, distribution is partly ensured by third-party, i.e. non-NSI and non-CESSDA organisations (e.g. Spain). Where NSI arrangements are rather friendly, moving towards a CESSDA-mediated system is likely to be less well-accepted unless an upgraded service is provided.

Regardless of these differences, a major gap for the full CESSDA membership is the total lack of an agreement between any of the Data Archives and Eurostat so far, largely due to the current European directive which allows Eurostat to contract only with research institutes.

In sum, CESSDA organisations play a role in the dissemination of government statistics, but strong differences across countries exist. They are due to a variety of factors, ranging from national legislation to the organisations' own history and culture. Where CESSDA members do not offer pathways for access to government data, NSIs (and sometimes, third-party institutions) do so, with mixed results. No CESSDA organisation offers any mediation services for Eurostat data.

4.1.4 A case for improved cooperation between CESSDA and Eurostat/NSIs

While some CESSDA organisations have established relationships with NSIs long ago, increasing privacy protection concerns and a surge in demand for administrative and combined administrative/survey datasets have led Eurostat and NSIs to manage confidentiality by setting up access arrangements for researchers under their own control. NSIs have also increased the part of anonymised data they distribute directly, thank to technical improvements (in particular, web-based tools) that enable to disseminate data freely

from their websites in the form of PUFs and tables. Hence, the amount of government data (both confidential and anonymised) that is disseminated directly by producers is increasing in most countries. Researchers are already using these facilities intensely and some of them have established direct, bilateral relationships with providers. This trend potentially challenges CESSDA as a single, comprehensive European system, and may jeopardize the efforts that some members have already made to secure a role in government data dissemination in their home countries. It also entails the risk of growing complexity and opacity for researchers, preventing the creation of a single point of contact for their data needs and increasing the time and effort they need to search for appropriate data sources. In this sense, the project to upgrade CESSDA calls for strengthened cooperation with NSIs, Eurostat, and other relevant public-sector data producers. This will also reassure those researchers who are already using NSI and Eurostat data facilities and may be unwilling to lose them for a new system of which they feel they have limited knowledge. The recent evolution of the legal framework offers an opportunity for CESSDA to play a role in the construction of a European system of access to confidential data in cooperation with Eurostat and NSIs, thereby extending the experience of the few countries where CESSDA members already participate in the process of dissemination. It is also anticipated that NSIs and Eurostat will be willing to share the rising financial burden of dissemination to researchers.

The current tendencies of NSIs and Eurostat to reinforce their role as direct distributors of the data they produce challenges the role of CESSDA as a central infrastructure providing pathways for access to research data. Unless a framework of cooperation between CESSDA and NSIs/Eurostat is established, this tendency may have adverse consequences for researchers who will have to face more than one point of contact for their data needs.

4.1.5 The possible role of an upgraded CESSDA in improving data circulation

A major problem is very limited circulation of data across borders, whether they are directly disseminated by producers or distributed by a CESSDA member archive. Indeed despite the principle of non-discrimination within the European Union, almost the same access conditions apply to European and non-European researchers. While some anonymised data files are allowed to cross borders (e.g. tables and PUFs that are freely downloadable from the World Wide Web), this is rarely the case for de facto anonymised data and for confidential data. It remains to be seen whether remote access connections might alleviate this problem in future, and one problem is to determine the legal framework that will apply to breaches of confidentiality (see also the CESSDA PPP tendered report on "Secure Remote Access system for an upgraded CESSDA RI"). To some extent, it may be possible to draw on the experience of non-CESSDA international databases such as LIS (Luxembourg Income Study), MTUS (Multi-national Time Use Study) and IPUMS (Integrated Public Use Microdata Series) which collect, harmonise and disseminate anonymised national government microdata in specific areas of interest. Either way, this issue calls for attention for the future CESSDA if it needs to develop from the current situation in which member organisations work with national resources, to create a common platform with a strong form of integration.

Despite recent improvements and a relatively favourable legislation, most microdata sets are not allowed to cross borders, even within Europe. No systematic response to this problem has been attempted so far. As a future European Research Infrastructure Consortium, CESSDA will need to tackle this problem and propose possible solutions.

4.2 Output of the Eurostat/ONS/CESSDA workshop

On December 3-4, 2008, the Eurostat/ONS/CESSDA workshop on "Microdata Access: New Developments and a Way Forward" took place. The workshop was meant to bring together Eurostat, NSIs, CESSDA Data Archives, and representatives of the European social science research community to discuss existing access arrangements and availability of official statistical data in European countries, so as to identify recent progresses made, best practices, and areas for possible improvement in view of the construction of the new European Research Infrastructure Consortium. Participants included the directors and staff members of the Methodology and Research direction of Eurostat, delegates from a large number of NSIs of European countries, almost all CESSDA Management Board members, and ten top-level representatives of researchers, invited by CESSDA. The agenda had been previously agreed by the three partners, and many of CESSDA's suggestions were accepted.

The workshop offered an excellent opportunity to discuss the potential and possibilities for CESSDA to act as data 'broker', mediating between public-sector data producers and data users not only at national level, but also at European level. The PPP was introduced in the first session by Kevin Schürer, who outlined results of the ESFRI (European Strategy Forum on Research Infrastructures) Roadmap, presented ongoing work, and hinted at future perspectives for the new infrastructure. The list of speakers also included other representatives of CESSDA. In particular, Roxane Silberman presented a paper on current accreditation arrangements throughout Europe, and Roxane Silberman and Paola Tubaro gave a presentation on access arrangements and the role of CESSDA Data Archives as intermediaries between government data producers and researchers. In addition, Vigdis Kvalheim illustrated the Norwegian experience of having a CESSDA Data Archive in charge of the dissemination of confidential government data. It is also worth mentioning an interesting presentation by Melanie Wright of UKDA on the new Secure Data Service, a pilot remote connection project that aims to offer researchers improved access to UK data, while still offering strong guarantees for data protection.

There was time for discussion in each session, and CESSDA delegates had an important opportunity to express their views. They participated actively in all debates, as did representatives of researchers. The workshop ended with a round table in which participants included Roxane Silberman for CESSDA and Bjørn Henrichsen for ESFRI, as well as representatives of Eurostat and ONS.

Afterwards, Eurostat prepared a webpage dedicated to the workshop (http://epp.eurostat.ec.europa.eu/portal/page?_pageid=2913,76672034&_dad=portal&_schema=PORTAL), where presentations have been posted; CESSDA's website has linked to it.

The conclusions of the workshop stressed current favourable trends, the interest of Eurostat for the role of CESSDA as a European infrastructure, and the crucial importance of building confidence between the different stakeholders. All participants including Eurostat insisted that this workshop can be regarded as a turning point in a new process, which might last long but has potential to lead to concrete follow-up steps in the short run. Possible actions were listed by participants (see below). Eurostat and some NSIs showed clear willingness to follow up on this process and foster further discussions for short-run and long-run cooperation.

For countries in which partnerships between NSIs and CESSDA organisations are still to be built, the workshop was an opportunity to share experiences and provide practical examples of how to proceed. Both CESSDA organisations and NSIs signalled a strong interest for possible future improvement. This first result highlights the potential of CESSDA to help extending good practices all over Europe.

Eurostat's willingness to cooperate with CESSDA was clearly demonstrated a few months later, when Eurostat invited Roxane Silberman to attend a workshop on remote access to microdata sets from official statistics for scientific purpose, which took place on 19 June 2009. The workshop, in which many NSIs participated, focused on the different systems of secure remote access to microdata. A number of objectives were set:

- Discussing existing remote access solutions;
- Sharing knowledge between Member States and identifying best practices;
- Enabling participants to learn from each other's experience;
- Prioritising ways to progress in providing remote access to microdata for researchers;
- Defining a relevant action plan and possible cooperative actions including CESSDA.

The discussion at the workshop focused on 2 main issues:

- Existing remote access solutions and possible re-applications in other countries;
- Possible setting of a European Statistical System infrastructure for remote access to confidential data.

In this perspective, Eurostat presented the (currently open) Commission FP7 call of 30 July 2009 on Social Science Data Archives and Remote Access to Official Statistics as an opportunity for cooperation with CESSDA. Eurostat also invited interested NSIs to join in negotiations in view of preparing a common proposal with CESSDA.

4.3 Recommendations

Building on these results, recommendations for CESSDA emerge. They can be divided into:

1. Recommendations for the medium/long run on the role of CESSDA as a provider of access to government microdata, and the relations between CESSDA and other stakeholders;
2. Actions to be taken in the short run to prepare future improvements.

4.3.1 Recommendations on CESSDA as a provider of access to government data

To become a comprehensive and truly pan-European Research Infrastructure Consortium, CESSDA must be able to offer mediation services for government microdata, including both anonymised and confidential microdata.

A. Thus, the first sub-set of recommendations is for CESSDA itself, which must strongly encourage member organisations to reinforce their role in this area where it already exists, and build it elsewhere, so as to create conditions for more equilibrated growth. Specifically:

   A1. *Significant involvement in government microdata dissemination should be a condition for membership* included in the CESSDA statutes and required of all national partners.

   A2. In particular, major *efforts to reach agreements with governmental data producers should be required of CESSDA organisations that do not currently offer any mediation services for government data* (or that do so to a very limited extent).

A3. To support efforts in this direction by Data Archives that have little or no experience in this area, CESSDA should *set up a permanent sub-committee or expert group in charge of providing assistance to members that need to prepare a first agreement with governmental statistics in their home countries*. This sub-committee should provide information on how agreements have been negotiated in the countries where they already exist, and advise on legal, institutional, ethical and organisational issues.

B. The second sub-set of recommendations concerns the relationships between CESSDA and a first category of external stakeholders, namely Eurostat and NSIs. Indeed for the new Infrastructure to become a central point of contact for researchers who need to access data, CESSDA must integrate with, and not overlook, the data facilities that governmental statistical agencies are setting up on their own. Specifically:

B1. CESSDA *should explicitly give a place in its statutes to Eurostat and to other NSIs* (and possibly, to other government microdata producers).

B2. To achieve this, CESSDA should *make a choice between different possible models of operation*, and negotiate with Eurostat and NSIs on this basis. Three solutions are conceivable at this stage:

- A first model simply consists in recognizing explicitly the NSIs with which relationships exist at national level, and inviting all others (also including Eurostat) to join the process as observers.
- The second model acknowledges Eurostat and NSIs as providers of part of a service in the new European Research Infrastructure Consortium. Put differently, the new Infrastructure would be a single umbrella including both CESSDA members and NSIs/Eurostat.
- A third model can be conceived as a mix of the first two.

B3. To inform its decision, *CESSDA should in the short run set up a team to explore more closely these three options*, in order to provide a motivated and detailed assessment of their anticipated advantages and shortcomings.

B4. *CESSDA should right away start negotiations with Eurostat and NSIs on very specific issues that have emerged at the Eurostat/ONS/CESSDA workshop as areas where rapid progress is possible*, even in countries that lack any experience of partnership between Data Archives and NSIs.

- A first set of actions require no change in the legislation. *CESSDA should consider providing services of documentation under DDI norms and format conversion* (e.g. to convert Eurostat files into SAS format, not currently available from the institution itself).
- A second set includes actions that require negotiations but no change in legislation. In particular, *CESSDA should propose to design and manage a European accreditation system*, through a transborder agreement allowing better transmission of knowledge about national research systems and institutions.
- A third set includes actions that concern primarily Eurostat rather than NSIs; as such, they do not require changes in legislation but might prepare further steps for which such changes will be necessary. CESSDA should

start negotiations with Eurostat in order to pass an agreement that would allow it to *cover costs of access for Eurostat datasets at national and (at a later stage) at European levels*. This will be most welcome by researchers and is likely to reinforce CESSDA's position even if Eurostat will still remain in charge of data dissemination. This might be started with a sub-group of member archives only, and enlarged afterwards.

B5. To foster long-term collaboration on these and related issues, *CESSDA should promote the transformation of the Eurostat/ONS/CESSDA event of 2008 into a permanent forum* that could take place regularly (possibly every two years). The costs of this forum will have to be estimated carefully.

In the long run, an important objective is to secure a formal place for CESSDA (and researchers) in forums for negotiations at European level. Membership in CEIES could be aimed at, even though this will require a change in European legislation. In the near future, improved relationships between CESSDA, Eurostat and NSIs could prepare the conditions for achieving the goal of progressive recognition of CESSDA as a partner, for instance with more systematic invitations to Eurostat workshops.

C. The third sub-set of recommendations concerns the relationships between CESSDA and another category of external stakeholders, namely researchers. They are the final users of the new infrastructure and it is important to ensure their support. Specifically:

C1. Representatives of *researchers should be members of the Scientific Council of CESSDA, and possibly of the sub-committee* that is in charge of providing advice to members that need to establish partnerships with governmental statistics.

4.3.2 Actions to be taken during the preparatory phase project

To implement the above recommendations, further work needs to be done shortly. This includes both pursuing negotiations with Eurostat and NSIs, and collecting further, more detailed information as a background for future agreements. In practice, the next few months should be used for the following actions:

A. The WP10 team within CESSDA PPP should be able to *prepare annexes on specific issues and countries*, to be used in future to prepare agreements.

B. The WP10 team should *explore in greater detail the position of CESSDA as a European Infrastructure with respect to international databases such as LIS, MTVS and IPUMS*.

C. To enhance negotiations with Eurostat, CESSDA should propose to *set up a common Eurostat/CESSDA working group to examine the issues that emerged at the workshop* (see point B4 above). WP 10 should be associated to this process.

D. One aspect of this immediate collaboration is the perspective opened by the EC FP7 call on Data archives and remote access to official statistics. A common proposal of CESSDA with some NSIs will provide concrete ground for discussions and collaboration on a crucial issue for the future ERIC.

E. As CNRS-RQ and UKDA are currently exploring the possibility to cover costs to distribute Eurostat's anonymised microdata at national level, *CESSDA should coordinate these initiatives while also involving additional partners in other countries*.

5. References

1. General

1.1 Publications, working papers, and paper presentations

- Als G. (1993) *Organization of Statistics in the Member Countries of the European Community.* Volume I: *Essays on the 12 National Statistical Institutes. Comparative Study.* Volume II: *Analytic Reports.* Luxembourg, Office for Official Publications of the European Communities, Eurostat - Theme 9 - Series E.

- Bruengger, H. (2004) "The relationship between the fundamental principle on confidentiality and population censuses: Statement from the UNECE Statistical Division," UN Symposium on Population and Housing Censuses, New York, September 13-14.

- Charpin, J.-M. (2007) « L'évolution du rôle de coordination de l'INS au niveau national dans un système statistique européen renforcé », *Courrier des statistiques,* n. 121-122.

- Dale, A. (2003) "Research Access to Microdata: an Attempt to Provide a Context", *IASSIST Winter Quarterly*, pp. 16-19.

- Desrosières, A. (1999) « La statistique publique dans les pays européens », *Courrier des Statistiques,* « La statistique publique dans les pays européens », special report, n. 91-92.

- Desrosières, A. (1999) « En visite dans un institut de statistique étranger », *Courrier des Statistiques,* « La statistique publique dans les pays européens », special report, n. 91-92.

- Díaz Muñoz, P. (2007) "Providing Microdata for Research Purposes in the European Union", paper presented at the 17th Statistical Days (Radenci, Slovenia, November 5-7).

- Doyle, P. et al. (Eds.) (2001) *Confidentiality, Disclosure and Data Access: Theory and Practical Applications for Statistical Agencies*, North-Holland.

- European Communities (1999) *R&D in Statistics: Academic and Official Statistics Cooperation,* Proceedings of the Seminar on *Academic and Official Statistics cooperation*, held in Bucharest in September 1998. Luxembourg, Office for Official Publications of the European Communities – Eurostat.

- European Communities (2003) Proceedings of the 19th CEIES Seminar on *Innovative Solutions in Providing Access to Microdata* (Lisbon, September 26-27, 2002), Luxembourg, Office for Official Publications of the European Communities –Eurostat, Theme 1, General Statistics.

- European Communities (2004) *Monographs of official statistics. Work session on statistical data confidentiality*, Part I, Part II, Part III (Luxembourg, 7 to 9 April 2003). Luxembourg, Office for Official Publications of the European Communities – Eurostat, Theme 1, General Statistics.

- European Communities (2006) *Monographs of official statistics. Work session on statistical data confidentiality*, (Geneva, 9 to 11 November 2005). European Commission and United Nations Economic Commission for Europe. Luxembourg, Office for Official Publications of the European Communities – Eurostat, Theme General and Regional Statistics.

- Eurostat Statistical Methodology Newsletter (2007), on "Access to Microdata in the EU: latest developments", vol. 1, n. 1, June.

- Eurostat and UNECE (2005) *Glossary on Statistical Disclosure Control*, developed by the Joint Work Session on Statistical Data Confidentiality. Available at: www.unece.org/stats/documents/ece/ces/ge.46/2005/wp.45.e.pdf.

- Gensbittel, M.-H., Riandey, B., Quentin, C. (2007) "Appariements sécurisés: statisticiens, ayez de l'audace! », *Courrier des statistiques*, n. 121-122, mai-décembre, pp. 49-58.

- Giessing, S., A. Harausz, A. Hundepool, H. Lindkvist, J. Longhurst, E. Schulte Nordholt, A. Smukavec, *Report on the CENEX-SDC inventory.* Available at http://neon.vb.cbs.nl/cenex/

- Glaude, M. (2007) « Enjeux et défis de la statistique européenne », *Courrier des statistiques,* n. 121-122.

- Goldmann, G. (2003) « Le Programme des Centres de données de recherche au Canada. Une politique globale de la recherche empirique destinée à fonder l'action publique sur la connaissance des faits », *Revue Internationale des Sciences Sociales*, 2003/3, N° 177, pp. 457-465.

- Hægeland, T. (2007) "Using Administrative Data for Research Purposes", paper presented at the 17th Statistical Days (Radenci, Slovenia, November 5-7, 2007).

- Hundepool, A., J. Domingo-Ferrer, L. Franconi, S. Giessing, R. Lenz, J. Longhurst, E. Schulte Nordholt, G. Seri, P.P. De Wolf, *Handbook on Statistical Disclosure Control.* Available at http://neon.vb.cbs.nl/cenex/

- Kraus, F. (1998) "Towards a Data Infrastructure for Socio-Economic Research on Europe: Improving Access to Official Microdata at the National and the European Level", *EURODATA Newsletter* n. 7, Spring.

- Kraus, F. (2004) "Data Protection and Access to Official Microdata for European Research", paper presented at the Network of Economic and Social Science Infrastructures in Europe (NESSIE), Roundtable 4, on "Data Protection and Data Access Regulations and Guidelines", held on November 26-27, 2004, Luxembourg. Available at http://www.nessie-essex.co.uk/default.asp

- Madinier, C., J.-L. Lhéritier, J.-P. Puig (2007) « La mise en place d'un Code de bonnes pratiques, pilier de la nouvelle gouvernance statistique européenne », *Courrier des statistiques,* n. 121-22.

- McCaa, R. and S. Ruggles (2002) "The Census in global perspective and the coming microdata revolution," *Scandinavian Population Studies*, no. 13, pp. 7-30.

- McCaa, R. and W.L. Thomas (2003) "Archiving Census Documentation and Microdata: Preserving Memory, Increasing Stakeholders", *Notas de Población* vol. XXIX, no.75, pp. 303-320.

- McCaa, R., S. Ruggles, M. Davern, T. Swenson, K. Mohan Palipudi (2006) "IPUMS-International High Precision Population Census Microdata Samples: Balancing the Privacy-Quality Tradeoff by Means of Restricted Access Extracts," *Privacy in Statistical Databases,* Berlin, Springer, pp. 375-382.

- McCaa, R., S. Ruggles, M. Sobek, and A. Esteve (2006) "Using integrated census microdata for evidence-based policy making: the IPUMS-International global initiative", *African Statistical Journal*, no. 2(May), pp. 83-100.

- Motohashi, K. (1998) "Institutional Arrangements for Access to Confidential Micro-Level Data in OECD Countries", OECD Science, Technology and Industry Working Papers, 1998/3, OECD Publishing. doi:10.1787/041104314765

- Museaux, J.-M., Peeters, M., Santos, M.J. (2008) « Legal, Political and Methodological Issues in Confidentiality in the European Statistical System", in J. Domingo-Ferrer and Y. Saygın (Eds.): *PSD 2008*, Springer, pp. 324–334.

- Muthmann, R. (2007) "Integration of Survey and Administrative Data", paper presented at the 17th Statistical Days (Radenci, Slovenia, November 5-7, 2007).

- OECD (2007) *OECD Principles and Guidelines for Access to Research Data from Public Funding*, OECD Publishing.

- OECD (Nadim Ahmad, Statistics Directorate) (2007) "OECD Project: Assessing the Feasibility of Micro-Data Access and Entrepreneurship ICE", Copenhagen, 22-23 February.

- Olsen, Ø. (2007) "Administrative Sources and Quality of Statistical Data", paper presented at the 17th Statistical Days (Radenci, Slovenia, November 5-7, 2007).

- Rivière, P. (2007) « Le programme statistique européen 2008 - 2012 », *Courrier des statistiques*, n. 121-122.

- Rouault, D. (2007) « L'accès aux micro-données et la gestion de la confidentialité dans quelques INS européens », *Courrier des statistiques*, n. 121-122.

- Schroeder, P. (2004) "Cherishing the Working Memory of Science: Towards International Guidelines for Access to Data from Public Funding", presentation at NESSIE Roundtable 4, November 26-27, Luxembourg**.** Available at http://www.nessie-essex.co.uk/default.asp

- Silberman, R. (2007) "Administrative data and access regimes for research purpose: old and new questions. The French case". Paper presented at the 17th Statistical Days (Radenci, Slovenia, November 5-7).

- Statistics Sweden (2003) *Access to microdata in the Nordic countries*. A joint report produced by Statistics Denmark, Statistics Finland, Statistics Iceland, Statistics Norway and Statistics Sweden.

- UNECE (2003) *Statistical Confidentiality and Access to Microdata,* proceedings of a Seminar, organized by the Conference of European Statisticians (CES) together with Statistics Sweden, held in Geneva on June 10-12, 2003; New York and Geneva, United Nations Publications.

- UNECE (2007a) *Managing Statistical Confidentiality and Microdata. Principles and Guidelines of Good Practice*, handbook prepared at the request of the Conference of European Statisticians (CES) – Task Force on Confidentiality and Microdata chaired by D. Trewin, New York and Geneva, United Nations Publications: http://www.unece.org/stats/publ.htm

- UNECE (2007b) *Register-based Statistics in the Nordic Countries. Review of Best Practices with focus on population and social statistics,* New York and Geneva, United Nations Publications.

- Verma, V. (2002) "Comparability of International Social Statistics", associated paper, NESSIE Roundtable 1 (Essex, September 12-13, 2002). Available at http://www.nessie-essex.co.uk/default.asp

1.2 Web sites

- CENEX-SDC: http://neon.vb.cbs.nl/cenex/

- CESSDA: http://www.cessda.org

- *Comparative Official Microdata*, European project led by F. Kraus and P. Flora from 1998 to 2001 (as a subproject of EuReporting): http://www.mzes.uni-mannheim.de/projekte/mikrodaten/comi.htm

- Eurostat: http://epp.eurostat.ec.europa.eu/portal/

- IPUMS (Integrated Public Use Microdata Series)-International, of the University of Minnesota: https://international.ipums.org/international/index.html

- NESSIE project: http://www.nessie-essex.co.uk/default.asp

- UNITED NATIONS STATISTICS DIVISION, Country profiles on official statistics: http://unstats.un.org/unsd/nsoprofiles/Default.aspx.

- UNECE's "Handbook of Official Statistics in the UNECE Region": http://www.unece.org/stats/handbook/

- The World Bank's "Country Statistical Information Database", accessible at: http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,contentMDK:20541648~menuPK:1164885~pagePK:64133150~piPK:64133175~theSitePK:239419,00.html

2. Country-by-country

This section provides, for each country, name and web address of the National Statistical Institute (NSI) and the CESSDA Data Archive (DA), together with a list of publications and other relevant web sites.

2.1 Austria

- NSI: Statistik Austria (Statistics Austria), http://www.statistik.at/

- DA: Wiener Institut für Sozialwissenschaftliche Dokumentation und Methodik (WISDOM), http://www.wisdom.at/Default.aspx

2.1.1 Publications

- Eichwald, R. (2001) "The Austrian System of Social Surveys". *EuReporting Working Paper No. 34.* (Mannheim: Mannheim Centre for European Social Research).

- Temples, M. (Statistics Austria) (2007) "SDC Micro: A New Flexible R-Package For The Generation of AnonymisAnonymised Microdata – Design Issues and New Methods", Paper presented at the Joint UNECE/Eurostat work session on statistical data confidentiality (Manchester, UK, 17-19 December 2007), WP n. 31.

2.2 Czech Republic

- NSI: Český statistický úřad (Czech Statistical Office, CZSO), http://www.czso.cz/csu/redakce.nsf/i/home

- DA: Sociologický datový archiv (SDA), http://archiv.soc.cas.cz/

2.2.1 Publications

- CZSO, *Dissemination Policy of the Czech Statistical Office*, March 2008.

2.3 Denmark

- NSI: Danmarks Statistik (Statistics Denmark), http://www.dst.dk/
- DA: Dansk Data Archiv (Danish Data Archives), DDA, http://www.dda.dk/

2.3.1 Publications

- Borchsenius, L. (2005) "New Developments in the Danish System for Access to Microdata", Paper submitted at the UNECE/Eurostat work session on statistical data confidentiality, (Geneva, 9-11 November), available at http://www.dst.dk/
- Møller, B., L. Solbjerghøj, and F. Kraus (2001) "Official statistics in Denmark: Socio-economic microdata for research". EuReporting Working Paper No. 24. (Mannheim: Mannheim Centre for European Social Research).
- Thygesen, L., O. Andersen and O. Schnor (Statistics Denmark) (2003), "The Danish System for Access to Microdata – From on-site to remote access", paper presented at the CAED conference, London, 15-16 September.

2.4 Estonia

- NSI: Eesti Statistika (Statistics Estonia), http://www.stat.ee/
- DA: Eesti Sotsiaalteaduslik Andmearhiiv (Estonian Social Science Data Archive), ESTA/ESSDA, http://psych.ut.ee/esta/

2.5 Finland

- NSI: Tilastokeskus (Statistics Finland), http://www.stat.fi/
- DA : Yhteiskuntatieteellinen tietoarkisto (Finnish Social Science Data Archive), (FSD), http://www.fsd.uta.fi/

2.5.1 Publications

- Lindqvist, M. (2001) "The Finnish System of Official Social Surveys". *EuReporting Working Paper No. 25*. (Mannheim: Mannheim Centre for European Social Research).
- Statistics Finland (2004) "Use of Registers and Administrative Data Sources for Statistical Purposes. Best Practices of Statistics Finland", Handbooks 45, Helsinki.
- Statistics Finland (2002) "Quality Guidelines for Official Statistics", Handbooks 43b, Helsinki.

2.6 France

- NSI : Institut National de la Statistique et des Etudes Economiques (INSEE), http://www.insee.fr

- DA: Réseau Quetelet, http://www.centre.quetelet.cnrs.fr/

2.6.1 Publications

- Blanc, M. and Desrosières, A. (2003) « Between Decentralization and Coordination: An Analysis of the Specificities of French Ministerial Statistical Offices", *Courrier des Statistiques,* English series, n. 9, pp. 3-19.

- Kieffer, A. (2001) "Le dispositif public d'enquêtes socio-économiques en France: une brève description". *EuReporting Working Paper No. 33*. (Mannheim: Mannheim Centre for European Social Research).

- INSEE (2001) Special report on « Le système français de statistique publique », *Courrier des Statistiques*, n. 98-99, available at : http://www.insee.fr/fr/ppp/publications/collect_som.asp?first_doc=21&simple=&coll=8&num_pub=&titre=&deb_mois=&deb_annee=&fin_mois=&fin_annee=&mot_cle=

  This special report includes the following papers :

  - Puig, J.-P. "Le système français de statistique publique";
  - Detape, Y. and Lacroix, J. « Le système français de statistique publique : cadre général » ;
  - Detape, Y. and Lacroix, J. « Le système français de statistique publique : l'INSEE » ;
  - Detape, Y. and Lacroix, J. « Le système français de statistique publique :les services statistiques ministériaux » ;
  - Desrosières, A. « L'équilibre entre coordination et autonomie du SSP : la spécificité du cas français ».

- Isnard, M. (2006) "Statistics and individual liberties: recent changes in French law," *Courrier des statistiques*, English series no.12, pp. 26-30.

- Silberman, R. (1999) « Les sciences sociales et leurs données », Report for the Ministry of Research.

2.6.2 Web sites

- CNIS (National Council of Statistical Information): http://www.cnis.fr/
- CNIL (Privacy Protection Authority): http://www.cnil.fr/
- CCDSHS (National Committee for Social Science Data): http://www.recherche.gouv.fr/comite/shumaines.htm
- DARES (Research and Statistics Department of the Ministry of Labour): http://www.travail-solidarite.gouv.fr/espaces/travail/
- DREES (Research and Statistics Department of the Ministry of Health): http://www.sante.gouv.fr/drees/index.htm
- INED (Institute for Demographic Studies): http://www.ined.fr/

2.7 Germany

- NSI : Statistisches Bundesamt Deutschland (Federal Statistical Office), http://www.destatis.de
- DA: GESIS (Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen), http://www.gesis.org/ZA/

2.7.1 Publications

- Allmendinger, J., Kohlmann, A. (2005) „Datenverfügbarkeit und Datenzugang am Forschungsdatenzentrum der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung“, Allgemeines Statistisches Archiv, vol. 89, n. 2, S. 159-182.
- Bohr, J. (2007) „Abschlussbericht - MISSY Nutzerstudie“, *ZUMA-Methodenbericht*, n. 1.
- Bohr, J., Janssen, A., Lengerer, A., Lüttinger, P., Schroedter, J., Wolf, C. (2007) „Verbesserung des Zugangs der Wissenschaft zu Mikrodaten. Pilotprojekt zum Aufbau eines Servicezentrums für Mikrodaten der GESIS bei ZUMA“, *ZUMA-Methodenbericht*, n. 5.
- Brandt, M., Lenz, R., and Rosemann, M. (2008) “AnonymisAnonymisation of Panel Enterprise Microdata – Survey of a German Project”, in J. Domingo-Ferrer and Y. Saygın (eds.), *PSD 2008*, Berlin and Heidelberg, Springer-Verlag, pp. 139–151.
- Esser, H., Grohmann, H., Müller, W., Schäffer, K.A. (1989) *Mikrozensus im Wandel. Untersuchungen und Empfehlungen zur inhaltlichen und methodischen Gestaltung*, Band 11 der Schriftreihe Forum der Bundesstatistik des statistischen Bundesamtes, Stuttgart, Metzler-Poeschel.
- Forschungsdatenzentrum der Statistischen Landesämter (2006) „Amtliche Mikrodaten für die Wirtschaft- und Sozialwissenschaften. Beiträge zu den Nutzerkonferenzen des FDZ der Statistischen Landesämter 2005“. (Contributions to the 2005 user conference of the Research Data Centre of the Statistical Offices of the Länder, in the social and economic sciences).
- Forschungsdatenzentrum der Statistischen Landesämter (2006) „Amtliche Mikrodaten für die Agrar- und Umweltwissenschaften. Beiträge zu den Nutzerkonferenzen des FDZ der Statistischen Landesämter 2005“ (Contributions to the 2005 user conference of the Research Data Centre of the Statistical Offices of the Länder, in the agrarian and environmental sciences).
- Haas, A. (2001) „Die IAB Regional-Stichprobe 1975-1997“, *ZA Information,* n. 48.
- Hessisches Statistisches Landesamt (2007) *Das Forschungsdatenzentrum der Statistischen Ämter der Länder*, Hessisches Statistisches Landesamt, Wiesbaden.
- Kohlmann, A. (2005) „The Research Data Centre of the Federal Employment Service in the Institute for Employment Research”, *Schmollers Jahrbuch*, n. 125, pp. 437-447.
- Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (2001) *Wege zu einer besseren informationellen Infrastruktur. Gutachten der vom Bundesministerium für Bildung und Forschung eingesetzten Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik.*
- Kraus, F. (2001) “The German System of Offical Social Surveys”. *EuReporting Working Paper No. 36*, Mannheim, Mannheim Centre for European Social Research.
- Lüttinger, P., Riede, T. (1997) „Der Mikrozensus. Amtliche Daten für die Sozialforschung“, *ZUMA-Nachrichten* 41, vol. 21, pp. 19 - 44.

- Lüttinger, P., Schimpl-Neimanns, B., Wirth, H. and Papastefanou, G. (2003) „Mikrodaten (German Microdata Lab): Das Servicezentrum für amtliche Mikrodaten bei ZUMA“, *ZUMA-Nachrichten* 52, Vol. 27, p. 153 – 172.

- Lüttinger, P., Schimpl-Neimanns, B., Wirth, H. and Papastefanou, G. (2004) „The German Microdata Lab at ZUMA: Services Provided to the Scientific Community”, *Schmollers Jahrbuch*, n. 124, pp. 455-467.

- Lüttinger, P., Wirth, H. (2004) „Regelungen der Weitergabe von amtlichen Mikrodaten an die Wissenschaft in den Bundesstatistikgesetzen seit 1953“, GML Arbeitspapier 01/4. Also published as „Zur Weitergabe von älteren amtlichen Mikrodaten an die Wissenschaft“, *Allgemeines Statistisches Archiv*, vol. 88, n. 4, 2004, pp. 473-486.

- Mïller, W., Blien, U., Knoche, P., Wirth, H. (1991) *Die faktische Anonymität von Mikrodaten*, Band 19 der Schriftreihe Forum der Bundesstatistik des statistischen Bundesamtes, Stuttgart, Metzler-Poeschel.

- Papastefanou, G. (1996) „Verbrauchsdaten der amtlichen Statistik: Forschungspotentiale zur Untersuchung des Wandels der Lebensführung in den neuen Bundesländern“, ZUMA-Arbeitsbericht 96/04.

- Rolf, G., Zwick, M., Wagner, G.G. (2008) *Fortschritte der informationellen Infrastruktur in Deutschland. Festschrift für Johann Hahlen zum 65. Geburtstag und Hans-Jürgen Krupp zum 75. Geburtstag*, Nomos Verlag.

- Statistisches Bundesamt (1998) *Survey of German Federal Statistics*, Statistisches Bundesamt, Wiesbaden.

- Wagner, G.G. (2002) “Innovative Solutions for the Provision of Access to Micro Data. The Case of the German Socio-Economic Panel Study (SOEP)”, associated paper, NESSIE Roundtable 1 on “Data for Comparative European Socio-Economic Research: Challenges, Opportunities and Constraints” (Essex, September 12-13, 2002). Available at http://www.nessie-essex.co.uk/default.asp

- Wirth, H. (1992) „Die faktische Anonymität von Mikrodaten: Ergebnisse und Konsequenzen eines Forschungsprojektes“, *ZUMA-Nachrichten* 30.

- Wirth, H. (2006) „Anonymisierung des Mikrozensuspanels im Kontext der Bereitstellung als Scientific-Use-File“, Arbeitspapier n. 11, Methodenverbund "Aufbereitung und Bereitstellung des Mikrozensus als Panelstichprobe".

- Wirth, H., Schimpl-Neimanns, B. (2004) „Anregungen zur konzeptionnellen Überarbeitung des CAMPUS-file Mikrozensus 1998“, German Microdata Lab, Arbeitspapier 3/2004.

- Zühlcke, S., Zwick, M., Scharnhorst, S., Wende, T. (2003) „Die Forschungsdatenzentren der Statistichen Ämter des Bundes und der Länder“, *Wirtschaft und Statistik*, Statistisches Bundesamt, n. 10.

- Zühlke, S., Christians, H. (2006) „Nutzung amtlicher Mikrodaten durch die Wissenschaft“, presentation at the conference *Unternehmen und Arbeitsmarkt in Bewegung -Was gewinnt die Forschung durch amtliche Mikrodaten?,* Wissenschaftszentrum Berlin für Sozialforschung.

- Zühlke, S., Hetke, U. (2002) „Datenbedarf der Wissenschaft. Ein Bericht des Forschungsdatenzentrums der statistischen Landesämter über die erste Nutzerbefragung“, Statistische Analysen und Studien NRW, Band 6.

2.7.2 Web sites

- Committee for Social and Economic Data: http://www.ratswd.de/index.html

- Research Data Centres of the Federal Statistical Office and of the Statistical Offices of the Länder: http://www.forschungsdatenzentrum.de/
- Research Data Centre of the Employment Agency (BA-IAB): http://fdz.iab.de/de.aspx
- Research Data Centre of the German Pension Fund: http://forschung.deutsche-rentenversicherung.de/ForschPortalWeb/contentAction.do?key=main_statistik&chmenu=ispvwNavEntriesByHierarchy482
- German Microdata Lab: http://www.gesis.org/Dauerbeobachtung/GML/index.htm
- Research Data Centre of the Development in Education Quality Centre: http://www.iqb.hu-berlin.de/arbbereiche/fdz
- IZA-IDSC: http://metadata.iza.org/home.php

2.8 Greece

- NSI: National Statistical Service of Greece, http://www.statistics.gr/
- DA: Greek Social Data Bank (EKKE – GSDB), http://www.gsdb.gr/

2.8.2 Web sites

- Bank of Greece: http://www.bankofgreece.gr
- Employment Observatory-Research Informatics S. A.: http://www.paep.org.gr
- Hellenic Migration Policy Institute: http://www.imepo.gr

2.9 Hungary

- NSI: Központi Statisztikai Hivatal (Hungarian Central Statistical Office), KSH-HCSO, http://portal.ksh.hu/portal/page?_pageid=37,115776&_dad=portal&_schema=PORTAL
- DA: Társadalomkutatási Intézet Zrt.(TARKI), Social Research Group, http://www.tarki.hu/en/

2.9.1 Publications

- HCSO (2007) "The evolution of Hungary's statistical system. Future rooted in the past", Country paper prepared for the UN Statistics Division (UNSD), Budapest.
- HCSO (2005) *Strategy 2005-2008*, Budapest.
- HCSO (2006) *Annual report on the strategy of the HCSO, 2005*, Budapest.
- HCSO (2007) *Annual report on the strategy of the HCSO, 2006*, Budapest.

2.9.2 Web sites

- ECOSTAT (Governmental Institute of Economic Research, affiliated to HCSO): http://www.ecostat.hu/english/
- DRI (Demographic Research Institute of HCSO and the Hungarian Academy of Science): http://www.demografia.hu/angol_nyito.html

2.10 Ireland

- NSI: Central Statistics Office, http://www.cso.ie/
- DA: Irish Social Science data Archive (ISSDA), http://www.ucd.ie/issda/

2.10.1 Publications

- CSO (2005) "Underpinning the Reputation of the CSO by maximising the Value and Utility of Data".
- McBride, J.P. (2001) "The Irish System of Social Surveys". *EuReporting Working Paper No. 29*. (Mannheim: Mannheim Centre for European Social Research).

2.11 Italy

- NSI: Istituto Nazionale di Statistica, http://www.istat.it/
- DA: Archivio Dati e Programmi per le Scienze Sociali (Data Archive for Social Sciences), ADPSS, http://www.sociologiadip.unimib.it/sociodata/

2.11.1 Publications

- Boeri, T., Pellizzari, M. (2003) "La deontologia di chi produce e detiene dati statistici: dalla possibilità alla certezza dell'accesso", *Rivista Statistica,* Special issue on "I microdati tra scienza e privacy", anno LXIII, n. 4, pp. 649-62.
- Hallu, R. (1999) « La statistique en Italie : complexe et unie, sous la houlette de l'ISTAT », *Courrier des Statistiques,* « La statistique publique dans les pays européens », special report, n. 91-92.
- Hallu, R. (1999) « Anagrafe della popolazione : un registre ancien, une ambition nouvelle », *Courrier des Statistiques,* « La statistique publique dans les pays européens », special report, n. 91-92.
- Ichino, A. (2003) "Banche dati solo sui giornali", *La Voce*, http://www.lavoce.info/articoli/pagina755.html.
- Ichino, A. (2003) "Le perplessità di un utilizzatore di dati di fronte al 'Codice di deontologia e buona condotta per il trattamento di dati personali per scopi statistici e scientifici'", *Rivista Statistica,* Special issue on "I microdati tra scienza e privacy", anno LXIII, n. 4, pp. 673-83.
- Ichino, A. (2004) "Ancora sul codice deontologico", *La Voce*, http://www.lavoce.info/articoli/pagina1087.html.

- Ichino, A. (2005) "L'ISTAT che vorremmo", *La Voce*, http://www.lavoce.info/articoli/pagina1510.html.
- Ichino, A. (2006) "Perché il dibattito politico prescinde dai dati", *La Voce*, http://www.lavoce.info/articoli/pagina2093.html.
- Ichino, A., Rossi, N. (2003) "Se la privacy non tutela la ricerca", *La Voce*, http://www.lavoce.info/articoli/pagina718.html.
- Peracchi, F. (2003) "Alcune osservazioni su protezione dei dati personali e analisi statistica", *Rivista Statistica,* Special issue on "I microdati tra scienza e privacy", anno LXIII, n. 4, pp. 693-696.
- Peracchi, F., Viviano, E. (2001) "The Italian System of Social Surveys". *EuReporting Working Paper No. 21*., Mannheim, Mannheim Centre for European Social Research.
- Schizzerotto, A. (2003) "Significati sociali e culturali del codice deontologico per il trattamento di dati personali a scopi scientifici e alcune sue implicazioni per la ricerca sociologica", *Rivista Statistica,* Special issue on "I microdati tra scienza e privacy", anno LXIII, n. 4, pp. 703-711.
- Trivellato, U. (2003) "Protezione dei dati personali e ricerca scientifica", *Rivista Statistica,* Special issue on "I microdati tra scienza e privacy", anno LXIII, n. 4, pp. 627-48.

2.11.2 Web sites

- SISTAN: http://www.sistan.it
- Privacy authority: http://www.garanteprivacy.it/
- Bank of Italy: http://www.bancaditalia.it
- IDAss (Italian Data Archive for the Social Sciences, http://www.idass.unitn.it/
- Laboratorio Riccardo Revelli of the University of Turin: http://www.laboratoriorevelli.it/whip/whip_datahouse.php.

2.12 Luxembourg

- NSI: Service Central de la Statistique et des Etudes Economiques, http://www.statec.public.lu/fr/index.html
- DA: Centre d'Etudes de Populations, de Pauvreté et de Politiques Socio Economiques / International Network for Studies in Technology, Environment, Alternatives, Development, CEPS/INSTEAD, http://www.ceps.lu/

2.12.1 Publications

- Jungblut, J.M. (2001) "The Luxemburgian System of Social Surveys". *EuReporting Working Paper No. 32*. (Mannheim: Mannheim Centre for European Social Research).

2.13 Netherlands

- NSI: Centraal Bureau voor de Statistiek (Statistics Netherlands), CBS, http://www.cbs.nl/nl-NL/default.htm
- DA: Data Archiving and Networked Services (DANS), http://www.dans.knaw.nl/en/

2.13.1 Publications

- Desrosières, A. (1999) « La statistique aux Pays-Bas : informatisation et intégration, un projet futuriste », *Courrier des statistiques,* « La statistique publique dans les pays européens », special report, n. 91-92.
- Groot, de W. and R. Dekker (2001) "The Dutch System of Official Social Surveys". *EuReporting Working Paper No. 30.* (Mannheim: Mannheim Centre for European Social Research).
- Hundepool, A. and de Wolf, P.P. (2006) "OnSite@Home: Remote Access at Statistics Netherlands", in EUROPEAN COMMUNITIES (2006), *Monographs of official statistics. Work session on statistical data confidentiality*, pp. 47-52.
- Viglino, L. (1999) « Argus, gardien du secret statistique », *Courrier des statistiques,* « La statistique publique dans les pays européens », special report, n. 91-92.

2.14 Norway

- NSI: Statistisk sentralbyrå (Statistics Norway), http://www.ssb.no/
- DA: Norsk samfunnsvitenskapelig datatjeneste (Norwegian Social Science Data Services), NSD, http://www.nsd.uib.no/nsd/english/index.html

2.14.1 Publications

- Hægeland, T. (2007) "Using Administrative Data for Research Purposes", paper presented at the 17th Statistical Days (Radenci, Slovenia, November 5-7, 2007).
- Henrichsen, B. and Tønder, J.-K. (2003) „Cooperation between researchers and national statistics institutes – the Norwegian model". Paper presented at Workshop on Microdata in Stockholm, 21 – 22 August 2003.
- Olsen, Ø. (2007) "Administrative Sources and Quality of Statistical Data", paper presented at the 17th Statistical Days (Radenci, Slovenia, November 5-7, 2007).
- Rodríguez, N.J.I. and Mellesdal, A. (NSD) "Practicalities Towards Access To Micro Data From Official Registers: The Norwegian Model".
- Støttrup Andersen, A. (2001) "The Norwegian System of Social Surveys". *EuReporting Working Paper No. 26*. (Mannheim: Mannheim Centre for European Social Research).

2.14.2 Web sites

- Act of 16 June 1989 No. 54 (the Statistics Act). Available at: http://www.ssb.no/english/about_ssb/statlaw/statlov_en.html

- Act of 14 April 2000 No. 31 (Personal Data Act): http://www.datatilsynet.no/upload/Dokumenter/regelverk/lov_forskrift/lov-20000414-031-eng.pdf

2.15 Romania

- NSI: Institutul National de Statistica (National Institute of Statistics), http://www.insse.ro/cms/rw/pages/index.en.do

- DA: Arhiva Romana de Date Sociale (Romanian Social Data Archive), RODA, http://www.roda.ro/

2.16 Slovenia

- NSI: Statistični urad Republike Slovenije (Statistical Office of the Republic of Slovenia), http://www.stat.si/

- DA: Arhiv Družboslovnih Podatkov (Social Science Data Archives), ADP, http://www.adp.fdv.uni-lj.si/

2.16.1 Publications

- Smrekar, T. "Access to microdata at SORS", paper presented at the 17th Statistical Days (Radenci, Slovenia, November 5-7, 2007).

2.17 Spain

- NSI: Instituto Nacional de Estadistica (INE), http://www.ine.es/

- DA: Archivio de Estudios Sociales (ARCES), http://www.cis.es/cis/opencms/FR/index.html

2.17.1 Publications

- Hallu, R. (2005) « La statistique en Espagne : nationale, mais aussi régionale », *Courrier des Statistiques*, n. 115, pp. 5-18.

- Hallu, R. (2005) « Padrón et Censo, bases de la démographie en Espagne », *Courrier des Statistiques*, n. 115, pp. 19-26.

- Jimenez-Martin, S. and J.M. Labeaga (2001*)* "The Spanish System of Social Surveys". *EuReporting Working Paper No. 22*. (Mannheim: Mannheim Centre for European Social Research).

2.18 Sweden

- NSI: Statistika centralbyrån (Statistics Sweden), SCB, http://www.scb.se
- DA: Svensk Samhällsvetenskaplig Datatjänst (Swedish Social Science Data Service), SSD, http://www.ssd.gu.se/index.php?lang=en

2.18.1 Publications

- Hjelm, C.G. (2006) "MONA - Microdata ON-Line access at Statistics Sweden", in EUROPEAN COMMUNITIES, *Monographs of official statistics. Work session on statistical data confidentiality*, pp. 21-28.
- Statistics Sweden (2001) *The future development of the Swedish register system,* R&D Report 2001:1.
- Vogel, J. (2001) "The Swedish System of Official Social Surveys". *EuReporting Working Paper No. 27.* (Mannheim: Mannheim Centre for European Social Research).

2.19 Switzerland

- NSI: Bundesamt für Statistik / Office fédéral de la statistique / Ufficio federale di statistica (Swiss Federal Statistical Office), http://www.bfs.admin.ch/bfs/portal/de/index.html
- DA: Fondation Suisse pour la Recherche en Sciences Sociales, FORS, http://www.fors.unil.ch/

2.19.1 Publications

- Ah, von T. (2001) "Socio-economic Surveys of the Swiss Federal Statistical Office". *EuReporting Working Paper No. 35.* (Mannheim: Mannheim Centre for European Social Research).

2.20 United Kingdom

- NSI: Office for National Statistics (ONS), http://www.statistics.gov.uk/default.asp
- DA: UK Data Archive, UKDA, http://www.data-archive.ac.uk/

2.20.1 Publications

- Alexandre, V. (1999) « La statistique au Royaume-Uni : You Can Count on Us - with Confidence », *Courrier des Statistiques,* n. 91-92.
- Ravalet, Ph. (1999) « Les mésaventures de l'Average Earnings Index », *Courrier des Statistiques,* n. 91-92.
- ESDS (Dennison, K.) (2008) "Guide to good practice: microdata handling and security", ESDS Access and Preservation.

- Dunnell, K. (2007) "The Evolution of the United Kingdom Statistical System", Paper submitted for the seminar on "Evolution of National Statistical Systems", UN, 23 February.

- Guy, A., A. Fiacco and F. Kraus (2001) "The System of Social Surveys in the United Kingdom". *EuReporting Working Paper No. 28*. (Mannheim: Mannheim Centre for European Social Research).

- Marsh, C., Dale, A., and Skinner, C. (1994) "Safe Data versus Safe Settings: Access to Microdata from the British Census", *International Statistical Review,* vol. 62, n. 1, pp. 35-53.

- National Statistics (2004) *National Statistics Code of Practice. Protocol on Data access and Confidentiality*, London.

- Jackson, P. (2008) "Routes to access: what affects the decisions that the ONS makes about research data access", paper presentation.

2.20.2 Web sites

- Web site of Northern Ireland Statistics and Research Agency: http://www.nisra.gov.uk/

- Web site of the Wales statistical office: http://www.wales.gov.uk/keypubstatisticsforwales/index.htm

- Web site of the Scotland statistical office : http://www.scotland.gov.uk/Topics/Statistics

- Census Programme: http://census.ac.uk/

- LES Research Laboratory Data Service : http://rlab.lse.ac.uk/itsupport/data/default.asp