Spatial Data Infrastructure for statistical production: challenges and opportunities

Helena CORDEIRO

Vice-President, Statistics Portugal, Portugal

Francisco VALA

Head of Unit – Territorial Statistics, Statistics Portugal, Portugal

Ana SANTOS

Head of Unit – Geoinformation, Statistics Portugal, Portugal

Abstract. The purpose of this paper is to promote a discussion on challenges and opportunities for the establishment of a Spatial Data Infrastructure (SDI) as an integral part of statistical production and dissemination. We argue that, besides a growing demand for territorial statistics and the very "natural" spatial dimension that statistical production presents, the establishment of a common SDI to support the production of European Statistics, based upon the coordination of input harmonization at European level, can contribute to an enhanced and more efficient production cycle throughout all its phases. In addition, we argue that it will also contribute to the effective implementation of the "Vision" in the ESS. The paper is organized as follows: first, we focus on the advantages of establishing a SDI to support the production of European Statistics throughout the European Statistical System (ESS); then, we address our experience at Statistics Portugal with SDI and its medium-term challenges for further development and the need of articulation with coordinated actions at the European level.

1. Introduction

The importance of space on statistical information can be assessed on the basis of different perspectives. First, we consider how the current social and economic context has added value to the territorial dimension. The current spatial complexity of some cross-cutting issues has reinforced the demand for territorial statistics by both the public administration and the civil society. Second, we argue that the role of space on statistical production process should be tackled in more than one dimension. In fact three dimensions are worthy addressing: an *organizing* dimension, a *context* dimension and as an *explanatory* variable.

Additionally, we find that the establishment of a common Spatial Data Infrastructure (SDI) for statistics purpose across Member States will contribute to the effective implementation of the

This paper elaborates on the presentation made by Statistics Portugal at the workshop on *Combination of geospatial information and statistics* held in Luxembourg, on 15-16 May 2012. Statistics Portugal wishes to acknowledge the comments and suggestions of the participants during the workshop. This version of the paper has also benefited from other presentations at the workshop.

"Vision" throughout the ESS (European Commission, 2009) by, simultaneously fostering multidimensional outputs and improving the efficiency of the ESS production. The search for data production efficiency, the growing scope, quality and use of registers, as well as the spreading of georeferenced data production, pose both a challenge and an opportunity to incorporate a wider spatial dimension on statistics production, including the sampling designs, and on dissemination.

Given the arguments in favour of the relevance of geographical data for statistics, the aim of this paper is to present the SDI and its role within statistical production, taking Statistics Portugal's experience and outputs as supporting examples. This paper is organized as follows: first, it focuses on the advantages of establishing a SDI to support European Statistical System (ESS) production; then, it addresses the experience of Statistics Portugal's SDI and its medium-term challenges for further development and the need to develop coordinated actions at the European level.

2. On the advantages of SDI within the ESS statistical production

2.1 The growing demand for territorial statistics

The current social and economic context, particularly marked by globalization and the relevance of environmental issues, has added value to the territorial dimension of social and economic reproduction processes. The current spatial complexity of these cross-cutting issues, given their local diversity and the different capabilities of territories to deal with them, has reinforced the knowledge of the importance of place-based approaches (Barca, 2009; Samecki, 2009) and increased the demand for territorial statistics by both the public administration and the civil society in order to support the implementation and monitoring of public policies, firms' assessment of the impact of territorial assets on market risk and also households' decisions regarding living and working conditions. This means that there is more awareness of the importance of not only keeping track of countries' performance, but also of the space of "everyday life" of both citizens and organizations.

The Local Agenda 21 settled by the United Nations in 1993, and its follow-up in the Rio+20 Conference on Sustainable Development, or the OECD's Local Economic and Employment Development Programme (LEED) are two examples that put in evidence the relevance of regional and local action to enable development and so the need to have more territorial statistics. In this context, it is worthwhile mentioning the OECD´s Territorial Development Policy Committee (TDPC) as a major international forum for discussion and exchange of experience in the field of regional policy, in particular regarding territorial development policies to promote regional competitiveness and effective and innovative governance, with these activities being directly

supported by a Working Party on Territorial Indicators (WPTI)[1]. Furthermore, at the European level, the Europe 2020 strategy coins a third dimension on cohesion approach – territorial cohesion – which will indeed require statistics to base its monitoring (European Commission, 2010). The European Commission is also suggesting that the next round of structural funds should be guided by a wider territorial perspective which clearly demands a sound set of territorial statistics (European Commission, 2012).

This growing demand for territorial statistics has been centred on the need i) to provide greater spatial disaggregation in traditional statistics, namely through NUTS level 3, Local Administrative Areas (LAU) or geographical grid datasets; ii) to make meaningful spatial units available in order to portrait the diversity of performances within a country – e.g. urban / rural areas, suburban areas / traditional centres, ultra peripheral regions / metropolitan regions, labour market areas –, and yet iii) to respond to new statistical demands with indicators that express territory-based phenomena – like data to support spatial planning, climate change and quality of life monitoring and the spatial organization of production.

Nevertheless, it is next argued that the relevance of geographic information on statistics goes well beyond this output of demand-driven perspective focused on the importance of a territorial dimension on statistics.

2.2 The three dimensions of space to be tackled by statistical production

The importance of geographical data can be summarized by three critical dimensions of space that must be tackled by statistical production and dissemination:

- an *organizing* dimension: space (like time) is the basis on which lies almost all data collection, storing, analysis and dissemination of official statistics. This *spatial nature* of statistical projects reinforces the use of geographical information to properly address the location element present in all phases of statistical production – from the design phase (to efficiently collect the raw data) till the dissemination stage (for example, using it to allow for a visual perception of data). Georeferenced statistical units including standardized address coding is therefore essential to tackle this dimension and contribute to a true harmonisation of official statistics within the ESS;
- a *context* dimension: space becomes meaningful for a particular data set on a specific territorial disaggregation and delimitation as the significance of the events measured do

[1] The document *Redefining urban: a new way to measure metropolitan areas in OECD countries* stands as an example of integration of GIS with other sources to produce new social, economic and environmental information at the local level, taking advantage of the potential of GIS data sources to monitor regional changes (OECD, 2012).

change accordingly to territorial arrangements used to portray statistical results. Geographers have been for long focused on this issue by discussing what they call *modifiable area unit problem* (MAUP), that is, the effects of scale and delimitation on empirical results. There is therefore a need to have proper territorial units to tabulate statistical results taking into account the nature of the phenomena measured and the aim of the results which are presented;

- an *explanatory* dimension: when space becomes itself statistical information as it conditions the phenomenon at hand. This perspective can be summarized by Tobler's first law of geography – all things are related but nearby things are more related than distant things – and its implications on the *distance effect* (physical or time-distance) and *spatial interaction* (flows and accessibility). While this explanatory dimension of space is already widely used to interpret statistical results, it should be further incorporated directly in statistical indicators design by combining different sources of geographical and statistical information, as territorial assets condition social and economic performances.

The relevance of these dimensions of space within statistical production for the different ESS organizations depends greatly on their position along the phases of data production cycle of European statistics. The still dominating "augmented stovepipe model" (European Commission, 2009: 6) positions member states on the data collection and data processing vertical / domain architecture, while Eurostat compiles the data coming from individual member states – eventually integrates information from the separate domains – and finally releases European aggregates. While Eurostat output oriented activities could take directly advantage from greater use of geographical information by exploring the *context* and *explanatory* dimensions of space within statistical production, a richer perspective throughout the whole ESS production, including the phases of the production cycle assured by member states, stresses the potential of the *organizing* dimension.

The fact that activities of the National Statistical Institutes (NSIs) cover all phases of data production poses a set of different opportunities for the ESS efficiency. In particular, data collection associated activities put in evidence the following opportunities:

- in the upstream of data collection, e.g. management of registers and sampling optimization;
- on data collection, e.g. collecting simultaneously both spatial and statistical data and running field work;
- in the downstream of data collection, e.g. to support processes of data linking and data matching, namely to small areas estimation.

The strategy should therefore go beyond an *output-oriented approach* and focus on setting *input harmonization* among Member States, seeking common infrastructure designs, geographical data themes and data tools and models. Such an approach would place the role of geographical information in ESS statistical production as a way of moving from a "stovepipe model" of production into "integrated parts of comprehensive production systems (the so-called data warehouse approach)" (European Commission, 2009). This perspective can contribute to foster the multidimensional output that is envisaged on the vision for reforming the production method of European statistics and simultaneously to support the architecture of production focused on microdata integration at the very beginning of the production cycle, increasing the advantage of the use of registers.

This line of thinking calls for a shift of the discussion from the advantages of geographical data for statistics to the advantages and challenges of implementing a Spatial Data Infrastructure to support statistical production of the ESS.

2.3 Moving the discussion from geographical data to Spatial Data Infrastructure

Moving the discussion from geographical data to Spatial Data Infrastructure (SDI) does not mean to leave behind the "what" debate – the ouput perspective – but to focus primarily on the "how" – the input perspective – and to consider that the notion of a SDI can largely contribute to the efficiency of the whole ESS production and simultaneously foster the multidimensional output that is envisaged in the 2009 Vision.

At this point, one should distinguish the several concepts at stake. **Geographical data** is *the data or information that identifies the geographic location of features and boundaries on Earth*. While the official production of such data in Europe is usually outside the sphere of competences of NSIs, statistical authorities already widely use geographical data, such as regions, administrative boundaries, census tracks, grids or buildings, to support statistical production and dissemination. The location of statistical units – e.g. firms, facilities, households – should therefore be interpreted as geographical data. Geographical data should be positioned as an element of a statistical information infrastructure. Additionally, **georeferenced statistics** – *the production of statistical aggregates positioned in space by means of geographical features* – is the core business of statistical authorities, even if limited to spatial units of large size, such as NUTS level 2 regions. The process of integrating geographical data and statistical data is performed by **Geographic Information Systems (GIS)**, which are already a common place within statistical authorities, and are usually described as *the tools used to gather, transform, manipulate, analyze, and produce information related to the surface of the Earth*. Nevertheless, there is a wide potential on using GIS

tools not only to produce new statistics (georeferenced or not) by integrating different sets of georeferenced statistics and those with geographical datasets (the output perspective), but also to support the organizations across the whole process of production (the input perspective).

In turn, the SDI concept, while embedded in GIS, it is a wider one. A standard concept of **SDI** could be described as *the set of resources, standards, technologies, policies, and legal, administrative and organizational frames required for the effective collection, use and sharing of geographical data in the statistical activity*. This means that, in comparison to GIS, SDI concept adds value to metadata as it favours the idea of sharing data and interoperability which, in turn, demands a well organized and standardized infrastructure. One can distinguish GIS and SDI from the problem perspective – the previous is project-oriented and the latter is infrastructural-oriented – and from the organizational point of view – GIS is intra-organizational oriented and SDI is inter-organizational oriented. SDI therefore covers a *human* dimension (producers and users, human capital), an *administrative* dimension (legal frame, organization, capacity building) and also a *technical* dimension (statistical data, standards and specifications, technology).

An enlarged usage of geographical data across the whole process of statistical production at the European level can only be achieved within the wider concept of SDI because it involves an inter-organizational dimension regarding cooperation between mapping agencies and national statistical offices, among statistical offices and between these and Eurostat.

The greatest challenge is therefore to implement a wise interoperability of the standard functions of a SDI with the current ESS statistical infrastructure so that they can become *integrated parts of a comprehensive production*. In fact, an integrated model of producing statistics will be able to cover integration across statistical domains at the level of NSIs and Eurostat, which means that the so-called cross-cutting issues or cross-domain concepts, such as environmental sustainability and quality of live, would be more effectively tackled. At the same time, an integrated model would promote the development of collaborative networks involving both the NSIs and Eurostat, creating opportunities for specialisation and harmonization by taking advantage of Member States best practices under Eurostat coordination and therefore promoting the *ESS production efficiency*.

2.4 A synthesis narrative on the advantages of SDI for statistical production and dissemination

In the previous section, it is argued that a SDI is a crucial concept in this process of enhancing the role of geographical data throughout the entire production cycle and thus promoting ESS efficiency by:

1) *speeding up*, that is, by promoting efficiency across the entire process of statistical

production – from sample optimization to data dissemination;

2) *deepening down*, in the sense that it allows for more spatial disaggregation which, while not necessarily an advantage itself, allows for more territorial arrangements and data integration, thus contributing to generate more meaningful statistical data and displaying it for relevant territorial units;

3) *adding meaning*, by integrating space into the conception of statistical indicators, combining different geographical and statistical datasets to produce new statistical information by means of spatial analysis; and,

4) *adding sight*, by providing an intuitive and appealing means to convey results through visual perception and GIS tools, and therefore a powerful tool to analyse and communicate results not only for dissemination purposes but also along the statistical production chain.

3. On the state of the art and medium term developments of Statistics Portugal's SDI

3.1 The state of the art of Spatial Data Infrastructure at Statistics Portugal

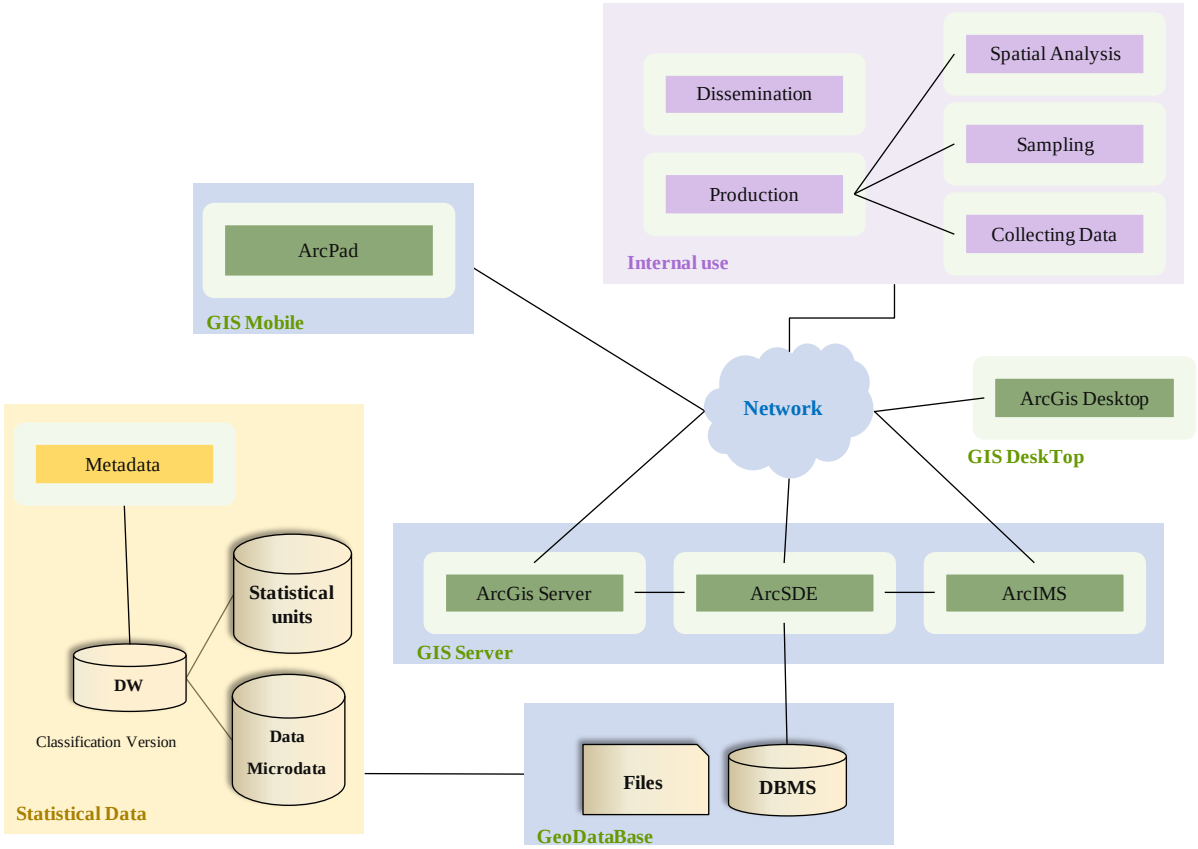
The use of cartography has supported census data collection at Statistics Portugal since 1981. In 1995, Statistics Portugal (SP) started the preparation of the cartographical infrastructure to support the 2001 census, which was named "Geographic Information Referencing Base" (BGRI 2001) and was based on Geographic Information Systems (GIS). Since 2006, with the production of the BGRI 2011 to implement the 2011 census, Statistics Portugal has been developing a SDI to support other statistical activities in a permanent effort to introduce the spatial perspective across the different phases of statistical production. At the same time, Statistics Portugal's SDI already provides ArcIMS map services to central bodies of public administration and standard WMS for general public, as the concretization of an INSPIRE commitment. SP provides, for example and among other data, the small statistical areas to the Ministry of Finance as an input for the creation and update of Municipal Taxes.

However, the *conceptual approach* of SDI in Statistics Portugal goes beyond the availability of a tool to support census operations and to build thematic maps. Statistics Portugal is now clearly aware of the importance of SDI to collect and handle spatial data, in order to support the statistical process and to manage the INSPIRE national thematic datasets under its responsibility.

Nowadays, Statistics Portugal's integrated system is designed to allow for: (1) an integration of the statistical data warehouse systems and the spatial data, including both statistical units that support statistical projects and the microdata obtained from statistical projects (the lower box in Figure 1); (2) the provision of the appropriate tools for statistical dissemination and to support statistical

production: the tools (and spatial data) to run spatial analysis, to design samples and to collect data (the box in the right top corner of Figure 1).

Figure 1. SDI's conceptual approach at Statistics Portugal



Regarding the *technological component*, Statistics Portugal has mostly used ESRI software for its desktop and server solutions (since 1997). More recently, open source software has made it easier to deal with costs constraints. So, at this point, one should stress the importance of moving from commercial software to open source software in this area too.

In what concerns the *spatial data component*, Statistics Portugal has developed its own geographical reference infrastructure (BGRI) which includes: (1) blocks with a two level structure – sections and subsections – integrated with the official administrative boundaries; (2) the road network; (3) residential buildings (under development to include other buildings, such as public facilities buildings and equipments); and, (4) the household register which uses the address as the location element.

Box 1. Introducing the spatial dimension to support households' statistical projects: the household register and sample optimization

At Statistics Portugal, the surveys directed to families/households have been made, since 1979, through a large master sample of dwellings – "Amostra-mãe" – extracted from the Census data. The difficulties in updating this master sample have lead to its depletion.

The 2011 Census micro data – for the first time georeferenced to buildings –, the access to different administrative sources and the implementation of the INSPIRE directive *Addresses* theme are all factors that together constitute an excellent opportunity towards a new strategy for the definition of a new sampling frame for family surveys at Statistics Portugal.

For this purpose, Statistics Portugal is already compiling a national file of addresses of all the dwellings (FNA), linked to the georeferenced buildings and to a street lines geodatabase.

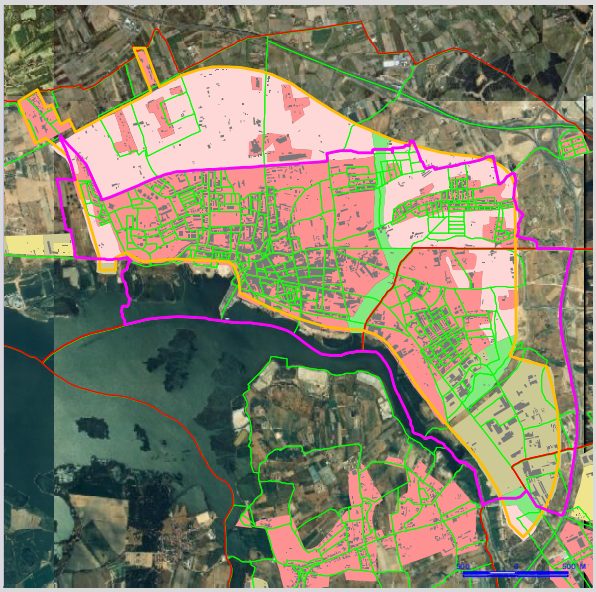
RUA	N PORTA	ANDAR	CP4	LADO	LUGAR	CP3	LOCAL POSTAL
RUA DOS DUQUES DE BRAGANCA	7	2	1200	ESQ	LISBOA	162	LISBOA
RUA DOS DUQUES DE BRAGANCA	7	3	1200	DT	LISBOA	162	LISBOA
RUA DOS DUQUES DE BRAGANCA	7	3	1200	ESQ	LISBOA	162	LISBOA
RUA DOS DUQUES DE BRAGANCA	7	<Null>	1200	<Null>	LISBOA	162	LISBOA
RUA DOS DUQUES DE BRAGANCA	7	SOBR	1200	ESQUERD	LISBOA	162	LISBOA
RUA DOS DUQUES DE BRAGANCA	7	RC	1200	DTO	LISBOA	162	LISBOA
RUA DOS DUQUES DE BRAGANCA	7	1	1200	<Null>	LISBOA	162	LISBOA
RUA DOS DUQUES DE BRAGANCA	7	2	1200	DTO	LISBOA	162	LISBOA

The FNA will be continuously updated using data from several administrative sources, like municipalities, tax revenue data, among others. The geographical support of the FNA creates an added value for sampling design, implementation and execution of the several statistical surveys conducted by Statistics Portugal.

Moreover, the BGRI is the base to build other geographies: some are directly linked to administrative boundaries – such as the NUTS classification, the degree of urbanisation (Eurostat DGURBA), the Urban Audit spatial units or the national urban / rural typology (TIPAU); others are derived exclusively from the BGRI like the statistical cities' boundaries (Box 2) or the Census localities.

Box 2. Creating relevant spatial units for statistical dissemination: the Portuguese statistical cities boundaries

Cities show a growing importance as driving forces of economic growth, competitiveness and employment. The Portuguese legal framework defines the conditions that allow a small town to achieve the city level. However, the specific legislation enacted by the Parliament when creating a new city does not account for the city territorial delimitation. In this context and based on the legislation that defines cities creation, Statistics Portugal has created the "statistical city" by defining the statistical criteria according to which it becomes possible to establish the territorial delimitation of Portuguese cities for statistical purposes.



In most cases, the definition of a statistical city relates to adjusting the urban perimeter provided by the legal instruments concerning programmed land use to the statistical subsections of the Information Reference Geographical Database (BGRI). Nevertheless, important infrastructures like industrial districts, commercial sites, ports and airports might also be included in the city limit even if located outside the urban perimeter. Such spatial delimitation allows for more data to become available for cities, at present mainly Census data. The further development of SDI and specifically the georeferencing process of statistical units will provide more data for cities.

The square grid as settled by INSPIRE can be directly linked with the buildings dataset once 2011 census final results are available. Statistics Portugal already was involved in GEOSTAT 1a project that defined the methodological approach for the production of a European population grid dataset. Statistics Portugal is currently involved in GEOSTAT 1b and will, in this context, produce the Portuguese population grid using the final 2011 census results.

Data component update and maintenance strategy is based upon:

- cooperation with INSPIRE organizations which produce / own geographical data. The institutions that are in the core of this strategy are (1) the national mapping agency that coordinates INSPIRE's implementing rules, acquires updated reference cartography of large scale and establishes the official administrative limits; (2) the municipalities which are strategic partners to update the buildings and addresses for INSPIRE themes; and (3) other organizations responsible for the transport network and the public services facilities themes.

- flow of data from municipal licensing construction system.

Box 3. Georeferenced data collection as part of construction statistics

SIOU is an information system effectively adapted to the municipal licensing construction system. It processes administrative information regarding:

– building permits,
– advance notifications,
– preliminary reports on buildings permits, and
– municipal projects.

This project is legally framed with the edicts of the law No. 555/99, Law 13/2000 and Decree Law 177/2001.

SIOU will be the main source for updating the Building's Geographical Base (BGE), as it results from the 2011 Building Census. At this moment, a new legal decision is being prepared to require municipalities to provide the X, Y coordinates of new buildings. The whole process of SIOU is being redesigned, including a new web application to submit the questionnaires. This application will allow the municipalities to access a module that identifies the X, Y coordinates of a new building. This module is specially designed for the municipalities that do not have a Geographic Information System already implemented.

3.2 The implementation of the INSPIRE Directive and other current applications

The INSPIRE Directive, which entered in force in May 2007, aims at establishing an infrastructure for spatial information in Europe to support policies or activities which may have an impact on the environment. INSPIRE is based on the infrastructures for spatial information established and operated by the 27 Member States of the European Union. The Directive addresses 34 spatial data themes needed for environmental applications, with key components specified through technical implementing rules. Statistics Portugal is involved in four out of eleven working groups and in five out of 34 INSPIRE themes: geographical names; addresses – which is central for the households register and to implement more data linking and data matching processes; statistical units; population distribution and demography; and, buildings. Additionally, Statistics Portugal is the leader of a working group (WG02).

Statistics Portugal association to the national INSPIRE working groups has evolved beyond its natural involvement in statistical units and population themes into a quite voluntary fashion, based on the relevance of Statistics Portugal data to a particular working group and our own strategic

guidelines on spatial data to support statistical activity.

Figure 2. INSPIRE Directive and Statistics Portugal (INE PT)

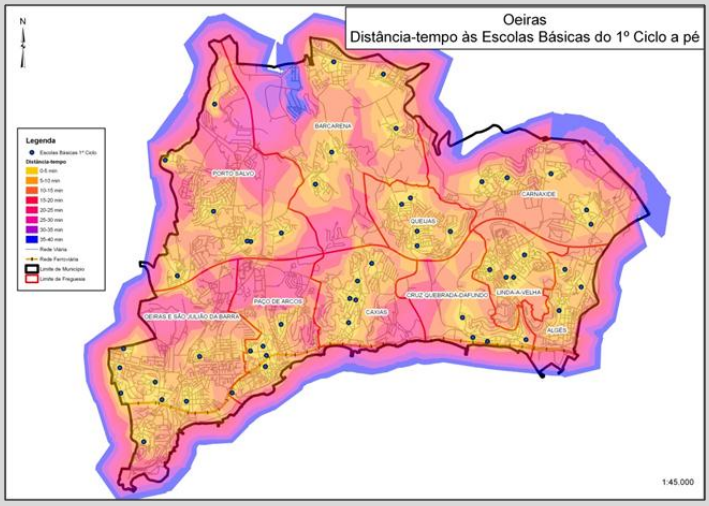
Working group	Annex / Theme	Formal responsible institutions
WG01	I.1 Coordinate reference systems	IGP, IGEOE, IH
	I.2 Geographical grid systems	IGP, IGEOE
	I.3 Geographical names	IGP, IGEOE, IH, **INE PT**, INAG, APA, CTT
	II.1 Elevation	IGP, IGEOE, IH, INAG
WG02	I.4 Administrative units	IGP, IGEOE, INAG
	I.5 Addresses	**INE PT**, CIM, CTT
	IIII.1 Statistical units	IGP, **INE PT**, CTT
	III.10 Population distribution and demography	**INE PT**
WG03	I.6 Cadastral parcels	IGP, DGCI, INAG
	III.2 Buildings	IGP, IGEOE, **INE PT**, IHRU, DGCI
Metadata and Services WG		APA, DGOT-DU,DRIGOT-RAM, INAG IP, IGeoE, IGP, IH, **INE PT**, LNEG I.P, SRCTE-RAA, ANMP (Câmara Municipal de Cascais)

The SDI already supports a variety of INSPIRE related projects at Statistics Portugal:

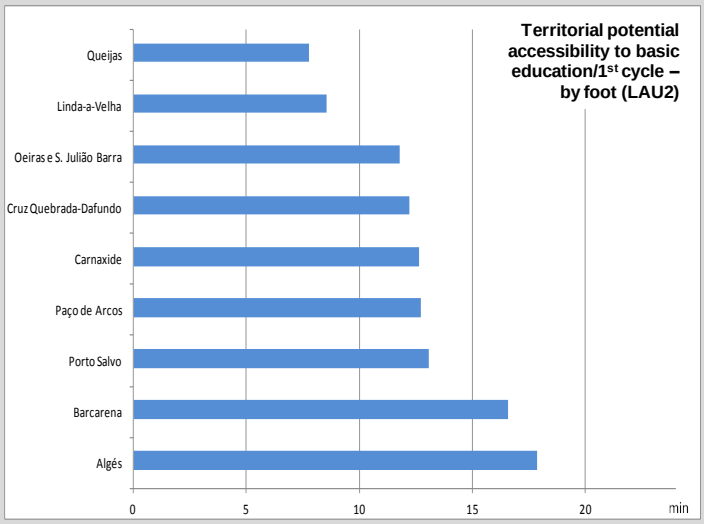
- geoinformation is present in official statistics portal through interactive thematic maps linked to a large number of statistical indicators, tools to explore the spatial units used for statistical production and dissemination and a facility to download census spatial units;
- in the statistical collection domain, supporting the web collection for municipal licensing (a pilot project that ran during the last few years and is under evaluation in order to be exhaustive in the coming year) and the update process of the dwellings register;
- in deriving statistical indicators by the combination of statistical and spatial data by means of spatial analysis: some with simple spatial analysis, others more complex like a pilot study on territorial and population accessibility to schools (Box 4).

Box 4. An experimental set of territorial accessibility indicators

Given the growing demand for statistical data related to mobility and accessibility, Statistics Portugal is developing an experimental set of territorial accessibility indicators as a result of the combination of already available statistical data. The combination of geographical data with statistical data is carried out through spatial analysis techniques with the future goal of regularly producing accessibility indicators.



The first tests were focused on the accessibility to schools and involved the following stages: (1) the selection of the teaching levels (the more advanced the level is the more distant the school is expected to be); (2) the delimitation of the areas under analysis (four regions were selected taking into account the population density and distribution); (3) the formulation of the accessibility indicators – establishing the starting and the ending points (place of residence and schools location), the concepts of time-distance and the travel means (walking or by car); (4) determining the most efficient time-distance routes according to transport means; and, (5) aggregating results to different statistical spatial units.



Further developments should focus on the robustness of the network model of analysis, on the use of other public facilities, and on the evaluation of the data quality of the results considering different input geographical data and scales according to the output statistical indicators.

3.3 The medium-term targets set by the 2013-2017 NSS programme

The general strategic guidelines for the official Portuguese statistical activity for 2013-2017, currently under consultation, stress the need to improve the efficiency of the statistical process through a methodological, scientific and technological development with a positive impact on the quality of statistics but it also mention the need to reinforce coordination and international cooperation. Specifically, it is argued that the focus should be on the use of administrative data, in particular those produced by public administrative bodies, in an attempt to rationalize the resources

available and to reduce respondent burden; on the full use of the geographic referenced infrastructure; on the reinforcement of spatial breakdown of statistical data; on the compiling and updating of National Statistical System (NSS) single files (of both statistical units and concepts), again favouring the rationalization of the resources and the quality of the output; and on the harmonization of procedures aiming at the integration of the various statistical domains.

The medium term-targets related with the SDI that are essential to the implementation of the above-mentioned guidelines will be:

- INSPIRE directive commitment activities, demanding an in-depth cooperation with other national organizations involved in INSPIRE and full compliance with INSPIRE implementing rules – focusing in particular on implementing normalized address coding in statistical projects and promote its use in registers and administrative data;
- to promote greater interoperability between spatial data and statistical data to support statistical production;
- updating and developing spatial units used for dissemination through the definition at the European level of national urban systems representative cities (to extend the urban universe covered by the Urban Audit project), updating the degree of urbanization classification based upon 2011 population grid data;
- at the national statistical system and based upon 2011 census blocks and census final results, redefining statistical cities boundaries, updating the urban / rural typology and evaluating in the Statistical Council the feasibility for implementing a Functional Urban Regions classification based on a labour market area approach;
- promoting spatial and statistical data integration to produce new statistical indicators by proceeding with the development of accessibility indicators to public facilities and of land use and urban expansion indicators and environmental indicators. The cooperation with other bodies of public administration responsible for managing administrative data related, for example, with the working force or social protection, plays a major role in this process of creating territorial statistics, especially in what concerns the already mentioned cross-cutting issues.

4. Final remarks

Statistics Portugal's experience suggests that the establishment of a common SDI to support ESS statistical production, under European coordination of input harmonization, can contribute to an improved efficiency of the production cycle throughout all its phases. Additionally, a strategy of

coordination will contribute to the effective implementation of the ESS's vision regarding the shift from a "stovepipe model" towards a "data warehouse approach" as put forward by the European Commission (2009). To achieve this ultimate goal of input harmonization, some further steps must be taken: for example, Member States consultations to assess the different stages of development regarding information technologies; the definition of guidelines regarding the degree of national NSIs INSPIRE's involvement; the carrying out of technical training actions and the sharing of relevant tools.

Side by side with coordination actions, there are some issues, both at the micro and at the macro level, which deserve careful attention and discussion. At the micro level, medium-term input harmonization should focus on production methods, multi-purposes statistics and opportunities for efficiency gains: to prioritize the phases of the statistical production cycle where spatial data will better foster the ESS efficiency; to identify statistical areas where spatial data is crucial to administrative data appropriation; to address quality criteria standards for geographical data used to produce new official statistics; to stress geographical metadata standards to be associated with statistical metadata; and, to define spatial data analysis procedures to be applied to families of indicators / themes.

At the macro level, medium-term input harmonization should focus on legal, economic and institutional challenges in implementing SDI concepts, components, systems and geographical technologies for statistical purposes, namely: barriers to and methods for achieving success in sharing geographical data; methods and models for planning, financing and implementing spatial data infrastructure or related initiatives; policies supporting creation of and responsible management of spatial data infrastructure resources.

This discussion should be framed by the principle of subsidiarity and by the scarcity of resources. In fact, it is clear that for achieving this outcome and collaboration to be intensified it is important to distinguish issues and processes which are candidates for co-operation from those that should prevail under separate responsibility and national competence. Secondly, the current reduction of resources at NSIs may affect the developments on the role of SDI in statistical production even if the discourse around it is an efficiency-based one.

5. References

[1] Barca, F. (2009), *An agenda for a reformed cohesion policy: A place-based approach to meeting European Union challenges and expectations*, Independent Report prepared at the request of Danuta Hübner, Commissioner for Regional Policy.

[2] European Commission (2009), *On the production method of EU statistics: a vision for the next decade*, Communication from the Commission to the European Parliament and the Council.

[3] European Commission (2010), *A strategy for smart, sustainable and inclusive growth*, COM (2010) 2020, Brussels.

[4] European Commission (2012), *The Programming Period 2014-2020: Monitoring and Evaluation of European Cohesion Policy – European Regional Development Fund, European Social Fund, Cohesion Fund: Guidance document on ex-ante evaluation December 2009*, Draft guidance document has been drawn up by the Directorates-General for Regional Policy and for Employment, Social Affairs and Inclusion.

[5] OECD (2012), *Redefining urban: a new way to measure metropolitan areas in OECD countries*, Note by the Secretariat, Statistics Directorate – Committee on Statistics (STD/CSTAT(2012)3).

[6] Samecki, Paweł (2009), *Orientation Paper on Future Cohesion Policy*, Communication from the European Commissioner in charge of Regional Policy.

[7] United Nations (1993), *Agenda 21: Earth Summit - The United Nations Programme of Action from Rio*.