

Spatial Data Sharing: A Pilot Study of French SDIs

Matthieu Noucher, Françoise Gourmelon, Pierre Gautreau, Jade Georis-Creuseveau, Adeline Maulpoix, Julie Pierson, Nathalie Pinède, Olivier Pissoat, Mathias Rouan

▶ **To cite this version:**

Matthieu Noucher, Françoise Gourmelon, Pierre Gautreau, Jade Georis-Creuseveau, Adeline Maulpoix, et al.. Spatial Data Sharing: A Pilot Study of French SDIs. ISPRS International Journal of Geo-Information, 2017, 6 (4), 10.3390/ijgi6040099 . halshs-01517140

HAL Id: halshs-01517140
https://shs.hal.science/halshs-01517140

Submitted on 2 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Open licence - etalab

International Journal of Geo-Information

MDPI

Article

Spatial Data Sharing: A Pilot Study of French SDIs

Matthieu Noucher [1,*], Françoise Gourmelon [2], Pierre Gautreau [3], Jade Georis-Creuseveau [1], Adeline Maulpoix [2], Julie Pierson [1], Nathalie Pinède [4], Olivier Pissoat [1] and Mathias Rouan [2]

1 CNRS, Lab. Passages, Maison des Suds, 12 Esplanade des Antilles, 33607 Pessac, France; Jade.Georis-Creuseveau@cnrs.fr (J.G.-C.); julie.pierson@cnrs.fr (J.P.); olivier.pissoat@cnrs.fr (O.P.)
2 CNRS, Lab. LETG-Brest, IUEM, rue Dumont d'Urville, 29280 Plouzané, France; francoise.gourmelon@univ-brest.fr (F.G.); adeline.maulpoix@univ-brest.fr (A.M.); mathias.rouan@univ-brest.fr (M.R.)
3 Lab. PRODIG, Université Paris 1 Panthéon Sorbonne, 2 rue Valette, 75005 Paris, France; pierre.gautreau@univ-paris1.fr
4 Lab. MICA, Université Bordeaux Montaigne, 10 esplanade des Antilles, 33607 Pessac, France; nathalie.pinede@u-bordeaux-montaigne.fr
* Correspondence: matthieu.noucher@cnrs.fr; Tel.: +33-5-56-838-406

Academic Editor: Wolfgang Kainz
Received: 27 January 2017; Accepted: 25 March 2017; Published: 28 March 2017

Abstract: Since the appearance of Spatial Data Infrastructure several years ago, there has been a tremendous increase in spatial data available on the Internet. This situation raises several research issues, in terms of identifying the content actually accessible via this medium and its impact on governance and local authority management. Our study proposes a mixed methodology applied to 45 French institutional infrastructures, to compare the objectives stated by their promoters, their content, and the actual services provided. The methodology, based on an analysis of interviews with Spatial Data Infrastructure (SDI) coordinators and their answers to questionnaires, as well as a study of their websites and an exploration of over 160,000 metadata in their metadata catalogues, produced varied results concerning data accessibility, stakeholder networks, the interoperability of tools, and informational equality in different regions. Despite the proactive stance of SDI promoters, only 15.7% of data are open-access. Their interoperability remains restricted to specific types of actors and themes. Although geocollaboration organised by SDIs is very active, it only concerns the public sector. These disparities also concern their informational dimension, as some regions have considerable resources at their disposal, but others do not.

Keywords: spatial data infrastructure (SDI); accessibility; interoperability; geocollaboration; informational equality

1. Introduction

Spatial Data Infrastructures (SDIs) are often seen as contributors to the vision of "Digital Earth", a multi-resolution, three-dimensional representation of the planet that makes it possible to find, visualize, and make sense of vast amounts of geo-referenced information on physical and social environments [1]. They are considered an essential driver of institutional data diffusion [2]. Due to the dual impact of system interoperability and changes in the legal framework, institutional spatial data are more and more accessible through the widespread deployment of SDI [3]. This facilitates the discovery, access, exchange, and sharing of geographic information and services among stakeholders from different levels in the spatial data community [4]. Standards are the key to developing interoperable platforms on the web. The promotion for access to geographic information aims to encourage its publication as open data, i.e., freely available to everyone to use and republish as they wish, without restrictions due to copyright, patents, or other control mechanisms.

Spatial data infrastructures are often reduced to simple websites for consulting mapping data on a region via a geoportal and downloading them from a geocatalogue, whereas, in fact, they are sociotechnical tools, revealing some of the major changes affecting the present system for producing and circulating spatial information. They are mainly run by government authorities, but deal with society as a whole, making it possible to observe the reorganisation of power relationships around data sharing on the Internet [5]. Thus, SDIs may be considered not as simple search engines but as real research subjects for exploring "informational governance" [6] in regions. Kooper et al. [7] define "information governance" as "*the set of activities aimed at establishing a normative foundation to facilitate and stimulate sense-making interactions*". By approaching them as interactive spaces where a network of actors can share and also co-produce data (geocollaboration), SDIs—and all of the associated working groups—may be understood as a key element for "sense making", thus participating in an analysis of informational governance. Furthermore, Mol [8] considers that "*information governance explains the emergence and working of nonhierarchical, nonlegal and noneconomic modes of governance, by focusing on the growing centrality of informational processes and resources*" (p. 21). A detailed analysis of a set of SDIs (and not one or two case studies) is required to discuss the non-hierarchical dimension highlighted by A. Mol.

Although publications on SDIs have been constantly on the increase for the past decade, Hendricks et al. [9] showed that the definition of SDI in the literature emphasized its technological components, thereby ignoring the major structural and human resource components of this tool. In a review of 237 significant, recent scientific articles [10], Coetzee and Wolff-Piggottalso emphasised that most research focused on technological developments and proofs of concept of their implementations. Considered as tools, SDIs are subject to technical analyses in areas such as system interoperability or data standardisation [11]. Organisational issues are essentially addressed to assess the reproducibility of governance systems from one country to another [12]. SDI evaluations are mostly based on quantitative criteria for purposes of comparison. The analysis of changes in data volumes and their similarities makes it possible to compare national SDIs and monitor them over time. However, data asset content is not yet being analysed. For example, the European Commission sees reporting as a benchmarking practice, aimed at creating a common cognitive framework, designed to make data comparable, without any consideration for content or relevance. A recent study contributes to this issue, by proposing a normalized evaluation of the performance, capacity, and availability of catalogue services [13].

Many scientific obstacles explain why these studies are so specialised. For instance, it is difficult: (a) to acquire a comprehensive view of all the information that flows through these decentralised platforms; (b) to identify SDI users, both producers or users of geographic information; and (c) to describe these uses and define them in relation to one another, since the use of geographic information amounts to a continuum of transformations as computerised data are so easily disseminated and so malleable. Indeed, digital spatial databases are malleable, i.e., sufficiently flexible to be shaped by a variety of processing operations (affecting their geometry, symbology, or semantic content) to suit the users' needs. Mashups that combine maps with data from various sources on a web page are good examples. Today, the development of Internet platforms providing access to spatial information contributes to the complexity of the "spatial information universe" [14]. However, understanding the way this universe functions and evolves is a major societal issue, as both public and private stakeholders, on their different levels, want to know the impact of these systems on governance and regional management. These considerations contribute fully to the current debate on critical data studies [15], concerning the contexts in which data, and spatial data in particular, likely to be transferred from one Internet infrastructure to another, are produced, analysed, circulated, and used.

Our study proposes a methodological framework for critical analysis of spatial data infrastructures, intended to complement and support this field of research, thus contributing to the necessary "task of conceptualisation on the basis of empirical research" mentioned by Kitchin and Lauriault [15] (p. 14). A mixed method capable of capturing large, diversified datasets (metadata,

website display, and results of the online survey and interviews) was applied to the French national and regional SDIs. The objective was to provide a comprehensive view of spatial data sharing, based on four dimensions. This necessitates a distinction between theory (what the promoters say) and practice (the reality for users) on *data accessibility*, the *interoperability of information systems*, *actor networking*, and *informational equality of territories*. *Data accessibility* was addressed by assessing the actual capacity of SDIs to provide access to reusable content by identifying properly-documented, downloadable data, suitable for use in a GIS. *Interoperability of information systems* concerned the capacity of SDIs to facilitate exchanges between expert systems and break down the silo approach of the 1990s (multiplication of isolated GISs for different specialities and regions with no interconnections). *Actor networking* was assessed by examining geocollaboration dynamics. Finally, we addressed the recurrent arguments justifying the implementation of these SDIs: *equality of information in the territories*. This implies that, unlike commercial systems, institutional platforms should provide homogeneous coverage of their area, in keeping with the principle of public service. This approach to the deconstruction of SDIs via the prisms of data, information systems, actors, and regions provides a critical analysis of the French context, comparing theoretical assertions with actual content. The reproducibility of this framework makes it possible to conduct analyses beyond the French context.

The paper is organised as follows. Section 2 presents the exploratory framework, the material, and the methods, while Section 3 presents the results for the four dimensions of spatial data sharing. These results are discussed and further work is outlined in Section 4.

2. Scope and Material

2.1. Material

In Europe, the aim of the INSPIRE Directive (2007/2/CE) is to deploy a Spatial Data Infrastructure in the European Community by relying on the SDIs of member States. In France, the National Council for Geographic Information (CNIG) ensures co-ordination between the national infrastructure (which has been deployed by the National Geographic Institute (IGN) and the French Geological Survey (BRGM) since 2005) and the myriad of thematic (soil, water, coast, etc.) or infra-national (departmental, regional, cross-border, etc.) infrastructures established since that time. Compared to other members state in the European Union, France is unusual in its large number of SDIs (Figure 1).

In 2012, the French association for geographic information (AFIGEO) identified 65 systems claiming to be SDIs [16]. This figure has grown constantly since 2008. The regional level alone represents half of the French infrastructure identified. Regional SDIs play an important role in maintaining links between local data and national standards, federating and coordinating them, as well as providing training [17]. We therefore chose to focus on national and regional infrastructures, which represented two thirds of the SDIs identified by AFIGEO. An update of the 2014 inventory enabled us to include 45 SDIs in our study: 16 national and 29 regional. The inventory was updated on the basis of exploration of the websites (checking that they were still active) and several interviews with coordinators of the national SDI network. Since the 2014 inventory, some SDIs had disappeared or merged and new ones had been set up, indicating the continuing instability of the institutional context.

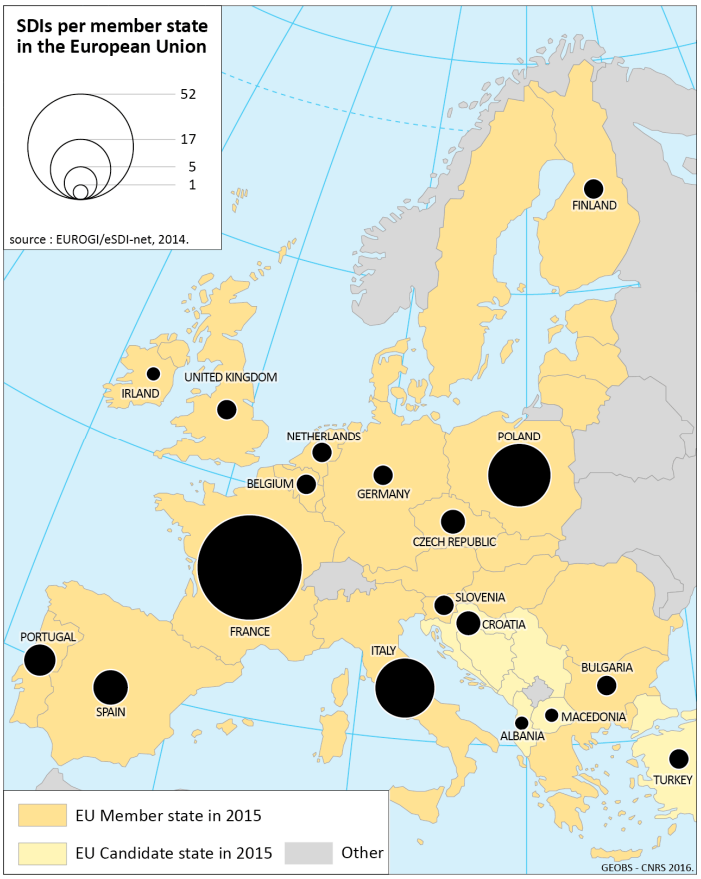


Figure 1. National and sub-national SDIs in European Union member states from 2009 to 2011 (source: EUROGI/eSDI-net 2014).

2.2. Methods

Ten research questions (Table 1) were presented in a mixed methodological framework [18] to analyse these 45 SDIs from the standpoint of data accessibility, tool and actor networking, and equality of information among regions. The qualitative approach was based on analysis of the editorial components of the websites and verbatim reports of interviews with several SDI promoters, to clarify their objectives and strategies. A quantitative approach, based on statistics and network analyses, was applied to data from a survey of the promoters of the 45 SDI and the metadata content of the metadata catalogues. The final element was a spatial analysis of the catalogue data. These diverse approaches were used to compare the promoters' assertions with the data actually circulated and analyse the effectiveness of the objectives of the SDIs.

Table 1. Four analysis topics explored via ten research questions.

Data Sharing Component	Research Question	Source
A. Data accessibility	**A1.** What do SDI promoters say about data accessibility?	Website and interviews
	A2. Are the reference data in SDIs downloadable?	Metadata catalogues
	A3. Are the data published in SDIs easily reusable?	Metadata catalogues
B. Tool interoperability	**B1.** What do SDI promoters say about interoperability?	Website and interviews
	B2. Are SDIs interconnected to accelerate data flows?	Survey
C. Actor networking	**C1.** What do the SDI promoters say about geocollaboration dynamics?	Website and interviews
	C2. Do SDIs promote geocollaboration dynamics?	Survey
	C3. Do SDIs facilitate the networking of diverse actors?	Metadata catalogues
D. Informational equality of territories	**D1.** What do SDI promoters say about informational equality?	Website and interviews
	D2. Do SDIs provide homogeneous coverage of their territories?	Metadata catalogues

2.2.1. Acquisition of Four Types of Data

Four types of data were gathered in the same corpus with the SDI as a central object. Assertion elements (from websites, survey, and interviews) and content elements (from metadata) were available for each SDI and were used for cross-sectional analysis:

- Websites

The editorial sections of the websites of the 45 SDIs in the study were explored. The pages presenting the system and its objectives were studied to identify the arguments justifying their implementation. Elements concerning the semantics, presentation, and organisation of content (e.g., graphics, hyperlinks, titles, and categorisation) were extracted and analysed.

- Survey

The survey was launched in May 2015 and lasted five months. This online survey of French SDI promoters was carried out using LimeSurvey survey manager. A series of closed questions was used to obtain an overview of the services (geoportal, geocatalogue, web services, etc.) and actions (training, communication, co-production, etc.) offered by each infrastructure. Open questions also invited the participants to clarify their objectives and targets. The promoters of 44 out of the 45 SDIs in our sample completed this survey.

- Interviews

A series interviews was conducted with 12 national and regional SDI promoters to obtain details of their strategy: objectives, targets, and future prospects. Semi-directive interview questions, concerning governance, organisation, and geocollaboration (co-design, co-production, and joint data control), were used to obtain information that was not accessible by automated metadata analysis. In addition to interviews in the six SDIs in two French regions (Nouvelle Aquitaine and French Guyana), six interviews were conducted on a national level with the sponsors of three SDIs, which, together, cover the entire country. The data harvested facilitated an in-depth analysis of the websites and survey answers, thus providing a more detailed overview of the SDI promoters' positions on the four issues studied.

- Metadata catalogues

The metadata from SDIs was produced using OGC CSW 2.0 web catalogue services. The contents of metadata tags considered useful for answering our research questions were stored in a database. A processing chain (Python) was implemented to automate web scraping (extraction from the CSW), storage in a PostGreSQL database, and analyses. It now takes little time to add a new metadata catalogue or update previously integrated metadata catalogues to reproduce analyses and work on the temporal evolution of the corpus. A description of this processing chain is included in Appendix A and the script, as well as the analyses using this method, are available in free-access (see Supplementary Materials). To check the reproducibility of this processing chain beyond French SDIs, we successfully experimented on SDIs from other countries (we processed 117 metadata from the SDIs of the Italian Province of Teramo) and global SDIs (we processed 5487 metadata available from the CSW of the SDI of IOOS® (Integrated Ocean Observing System)). For this article, the operational web cataloguing services of 37 of the 45 French SDIs were used to analyse 160,603 metadata sheets. Data extraction took place in summer 2016, between 22 June and 22 August. The databases were analysed using different tools, as required.

2.2.2. Data Analysis

The four types of data harvested were analysed in different ways (Figure 2).

- Verbatim quotes from websites and interviews and lexical analysis

Arguments identified during the analysis of websites and interviews led to the collection of a series of verbatim quotes that were used directly, without statistical processing, to retain the variety of opinions. A speech analysis, focusing on terminological occurrences, was also conducted to reveal the main categorical trends.

- Statistical processing of survey answers and metadata

Univariate statistical analysis, using R software, produced summary data to facilitate the use of survey results and exploration of the metadata sheets.

- Network analyses to process survey answers and metadata

Statistical processing of survey answers and the database, integrating all the content of the metadata tags, were completed by network analyses. In particular, connections between SDI were revealed by graphs based on geocatalogue harvesting.

- Spatial analysis applied to metadata

This was done by extracting 4 coordinates from the bounding rectangles. This extraction was based on the retrieval of the XML tags corresponding to the bounding (<EX_GeographicBoundingBox>) in order to subsequently generate data in GeoJSON format. PostGIS and Python processing are then used to produce a geographical data density map. This makes it possible to highlight the spatial patterns in the composition of local and national institutional geographical data assets. A description of this processing chain is included in Appendix B and the script, as well as the analyses using this method, are available in free-access (see Supplementary Materials). To check the reproducibility of this processing chain beyond French SDIs, we successfully experimented on SDIs from other countries (GeoBolivia from Bolivia and INDIA from Brazil).

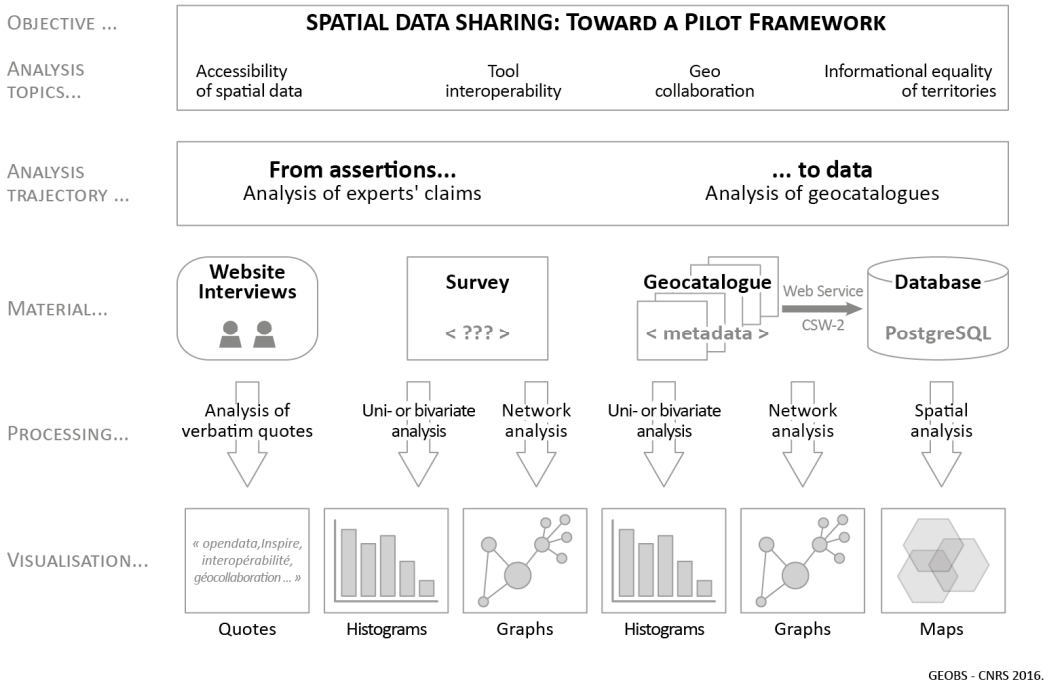


Figure 2. Analysis framework workflow.

Surveys and interviews with coordinators, as well as exploration of the websites and geocatalogues, thus constituted the backbone of the methodological framework implemented to assess the effectiveness of the SDIs' objectives (Figure 3).

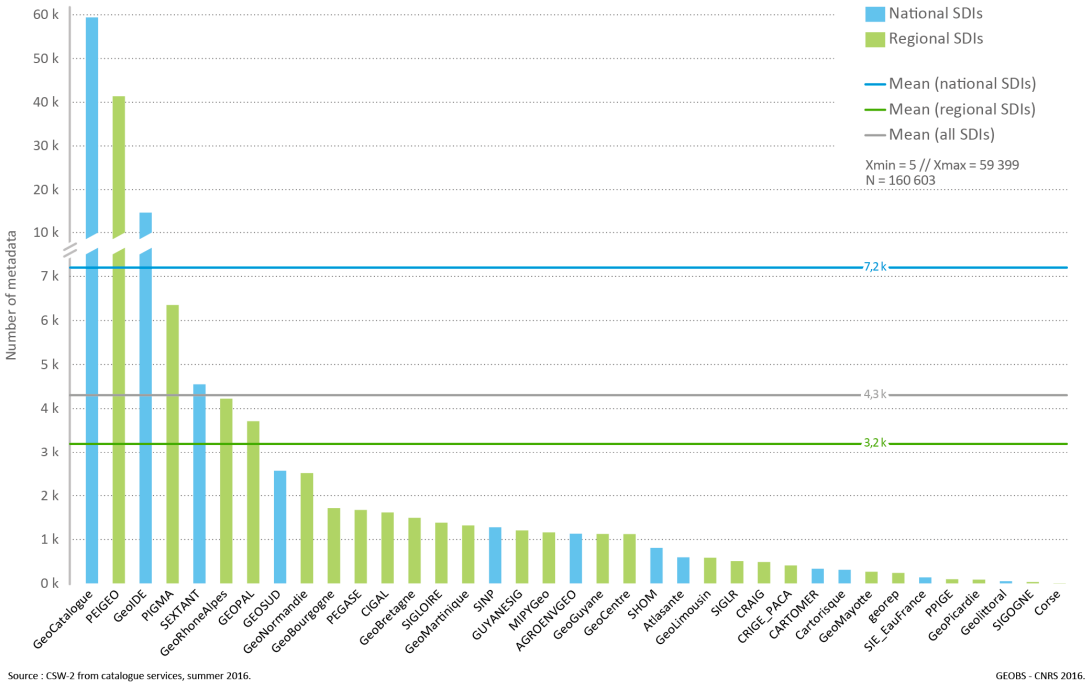


Figure 3. Volume of metadata from CSW of regional and national SDIs in France.

3. Results

All of the analyses applied to the four types of data (websites, interviews, survey, and metadata) gave results in the four components of spatial data sharing (Table 1 and Figure 2).

3.1. Data Accessibility

3.1.1. What Do SDI Promoters Say about Data Accessibility?

One third of the 45 regional and national SDI coordinators included in the study referred to the European INSPIRE directive in describing the objectives of their system. Of the 45 SDIs studied, 41 had websites with editorial pages, 35 of (78%) which included direct references to INSPIRE on their presentation page (some even devoted entire sections to explaining the history and content of the Directive). The origin of most of these SDIs, therefore, apparently follows the dynamic initiated by this legal framework. Paragraph (3) of the preamble to the Directive emphasises the importance of the accessibility and sharing of spatial information to justify its implementation. The issue of data accessibility is, therefore, at the heart of these dynamics and one of the priority objectives is to promote the wider use and reuse of spatial information.

3.1.2. Are the Data in SDIs Downloadable?

SDI websites were analysed to determine the volume of spatial data documented in geocatalogues (Figure 3). In addition to the national geocatalogue, which includes 59,399 metadata sheets, or a little over one third of the material (160,603), the 37 other catalogues contained an average of 4300 metadata sheets: an average of 7200 in the national catalogues and 3200 in the regional ones.

However, the vast majority of the documented data are not freely accessible. Thus, metadata searches including the key words "open data", to identify freely-downloadable data, moderate the idea of a "data deluge" likely to drown users! Only 25 of the 37 geocatalogues with operational CSW offered free-access to data and, on average, only 15.7% of their catalogue data are currently downloadable (Figure 4). A gap between the national and regional levels may be noted, with more than twice as high a percentage (19.1%) of data tagged "open data" in regional compared to national catalogues (8.5%).

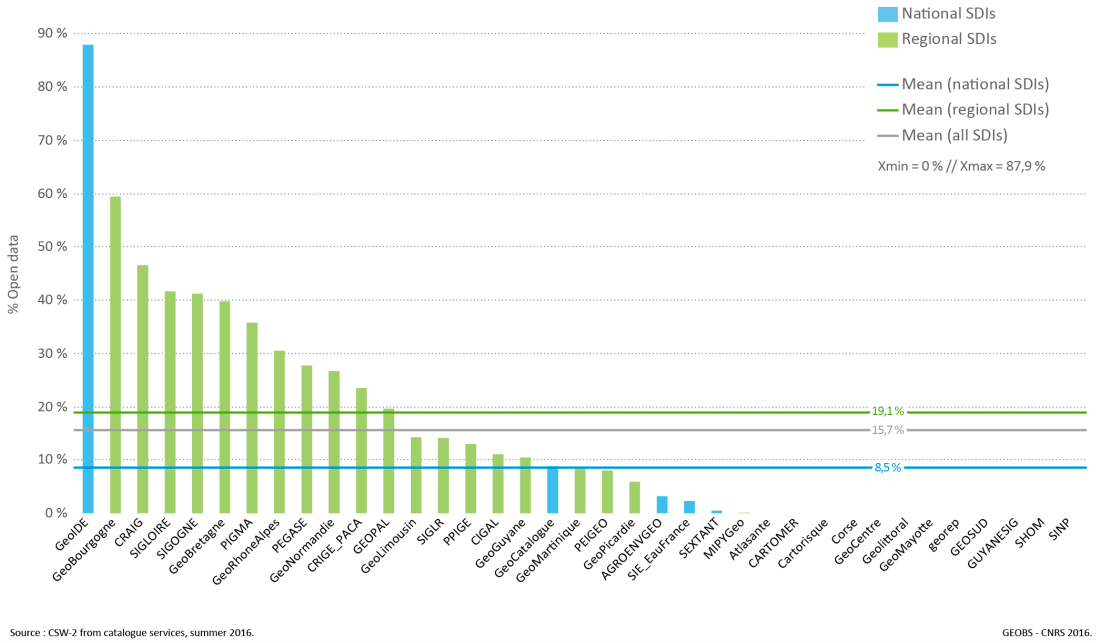


Figure 4. Percentage of open data in regional/national SDI geocatalogues in France.

3.1.3. Are the Data Published in SDIs Easily Reusable?

While our previous analyses provided a more nuanced picture than media announcements on the systematic opening of spatial databases associated with French SDIs, it is still true that data accessibility has increased considerably in recent years [15]. However, the actual possibility of reusing the data depends on compliance with documentation rules. Indeed, the metadata are indicative of the data processing conditions, providing, for example, information on validity space and time scales. Despite the use of standardised tools compliant with international standards (ISO 19115), our analysis of metadata fields revealed that content was often very imprecise, which hampered rigorous reuse of the data. Thus, we attempted to analyse several metadata fields to identify the sources and production methods of the spatial data presented. A textual analysis of the "genealogy" tag (supposed to provide a concise description of the history of the data set by presenting the resources, procedures, and processing implemented during their acquisition) revealed that this free-input metadata field was inoperable for systematic processing. The information was heterogeneous, referring to very different database or producer denominations, and was even, in many cases, non-existent. The study of time data tags provided another example of the limitations of metadata analysis. Thus, 95% of these data did not provide any information on the time depth of the data and the various "date" fields were completed in different formats (Figure 5).

Our results show some limitations in comparing SDIs with one another in providing differentiated interpretations of the standards. Even though all SDIs refer to the INSPIRE directive and the ISO 19115 standard, geocatalogues are still being rationalised and standardised. We felt that obstacles to analysing metadata fields concerning data genealogy and dating were significant indicators of two major, related difficulties: the low quality of the metadata input and the complexity of current standards (with many subtleties, like date formats, which are not understood by data producers, whose core business is rarely geomatics). However, this observation does not prejudge the quality of the data to be analysed in terms of their specification (internal quality) and uses (external quality or "fitness for purpose") [19]. This area of analysis is outside the scope of our study, which focused on a systematic exploration of metadata to visualise the contents of SDIs, without any distinction of quality.

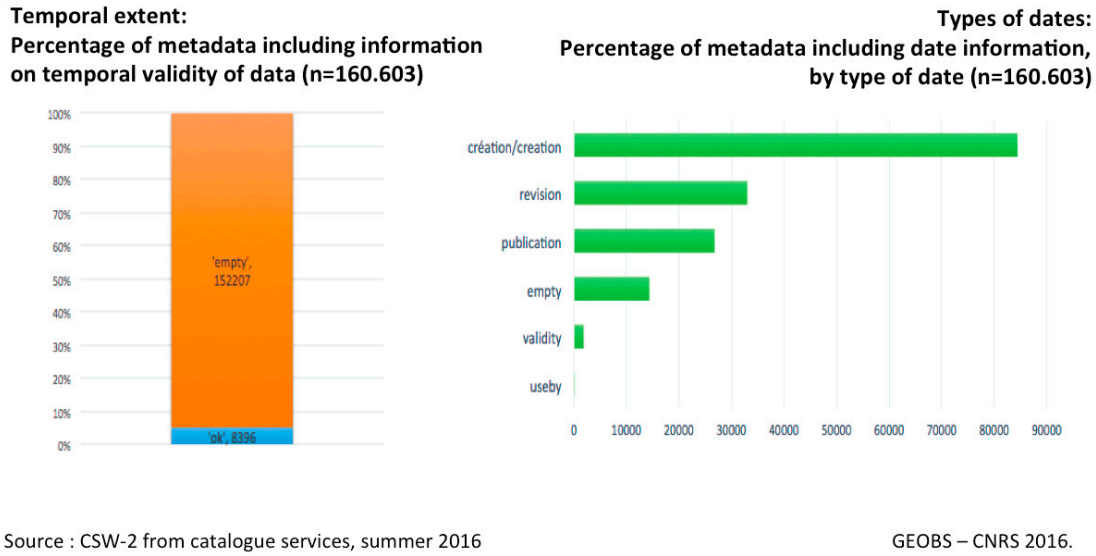


Figure 5. Time information in metadata sheets.

3.2. Tool Interoperability

3.2.1. What Do SDI Promoters Say about Interoperability?

One of the recurring arguments in favour of implementing SDI concerns these platforms' technical capacities to connect databases stored on distributed servers. Thus, 40 out of the 45 SDI promoters surveyed reported that they used web services to access remote metadata (CSW) or data (WMS, WFS) potentially distributed on several servers. These interconnection issues are at the heart of national spatial data infrastructure implementation strategy. According to answers to our survey, the on-going reorganisation (2015–2018) of the geoportal/geocatalogue (known as "geoportal v4") is being implemented "in the context of the emergence of the platform-state" and is aimed, in particular, at reinforcing "the mechanisms for aggregating regionally- or locally-produced data on a national level", to develop "chained geoprocessing with heterogeneous sources" or even to "improve harvesting/caching mechanisms". The platform-state is a national strategy to promote the emergence of new digital public services, based on improved data circulation. This concern with improving the circulation of information is also perceptible on a regional level, with SDIs that set themselves the objective of "developing new functions (web services) to facilitate connections in local authorities" or more generally "promoting the interoperability of different systems to reduce public spending and save time". By bringing together actors from the same local authority or working on the same topic, SDIs are supposed to decompartmentalize GIS that have been implemented in all business areas and on every possible scale since the 1980s. However, an analysis of their geocatalogues sheds some doubt on these claims.

3.2.2. Are SDIs Interconnected to Accelerate Data Flows?

Geocatalogue harvesting was analysed to assess the degree of interconnection of SDIs. Harvesting is a mechanism for collecting metadata from a remote catalogue and storing them on a local node to provide faster access. Harvesting is more than importing data, as the local and remote catalogue metadata are synchronised. For example, it is possible to access all the metadata in "departmental" catalogues and a selection of metadata from a national catalogue from a regional catalogue. Analysing the implementation of harvesting, therefore, makes it possible to identify links between SDIs. Furthermore, the harvesting process was studied by network analysis to identify the central (or isolated) nodes and any clusters (by displaying sub-networks). In addition to connections between SDIs, the survey obtained the first census (on a self-reporting basis) of connections between SDI and

observatory geocatalogues, open data portals, infra-regional infrastructures, etc. Thus, Figure 6 shows all the links between the SDIs in our study and Figure 7 integrates all the 2nd-level links, i.e., all the systems directly in contact with those SDIs.

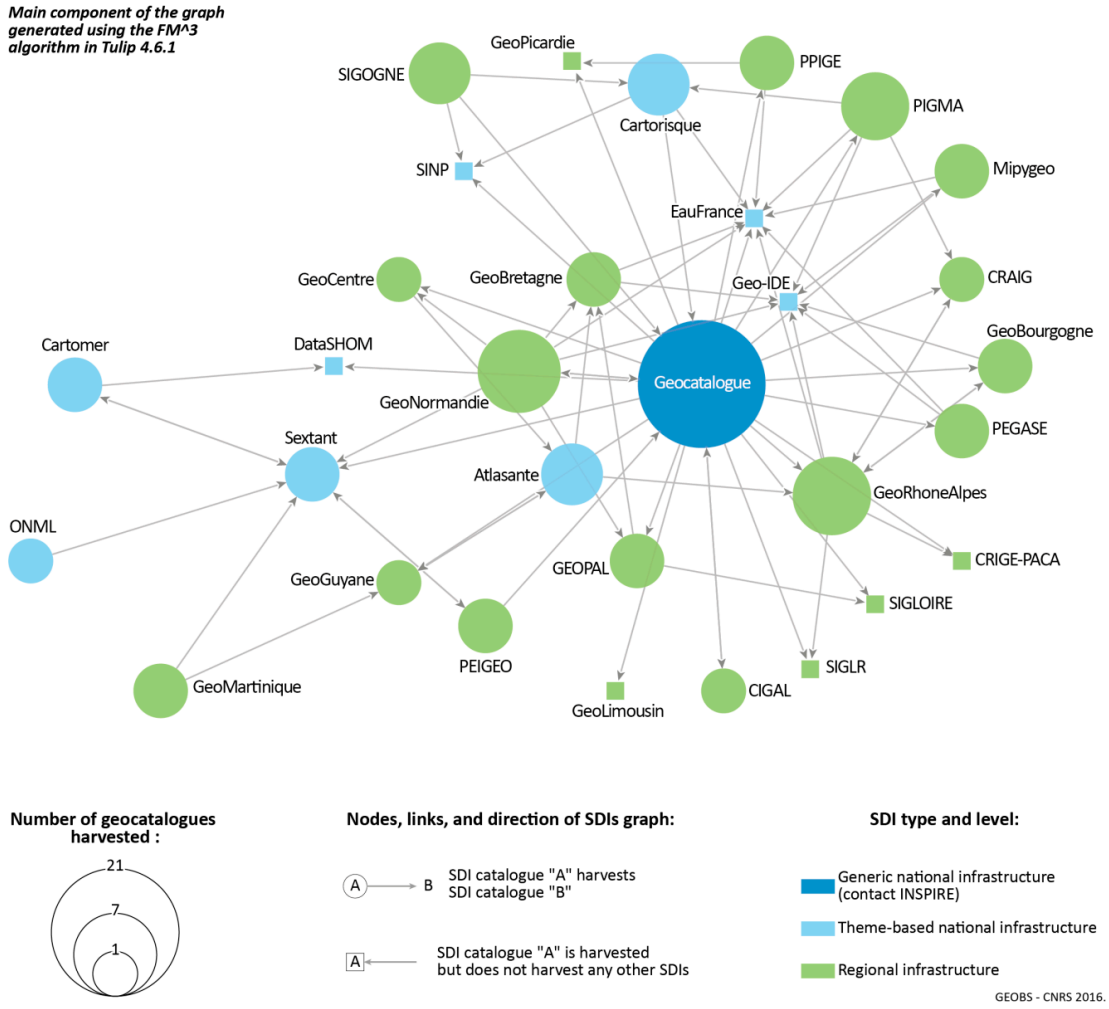


Figure 6. Network of inter-SDI links, established by analysing harvesting between the platforms included in this study.

The graph in Figure 6 shows the network of interconnected regional and national SDIs, highlighting the central position of the Geocatalogue, which constitutes the national focal point for reporting under the European INSPIRE directive and is unusual in having connections to 21 SDIs—four times the general average. This graph also shows several specific circuits for accessing information:

- Effects of scale are very marked in the centre-periphery structure of the graph. Most national SDIs are in the centre and the first outer circle gives a panorama of regional SDIs.
- Some infrastructures that make little use of harvesting techniques (small nodes) are centrally located as they are frequently harvested. This is the case of EauFrance and Geo-IDE.
- Theme-based network components may also be highlighted, for example, by connecting SDIs on maritime or coastal themes and in French overseas regions (Sextant—GeoMartinique or PEIGEO link).

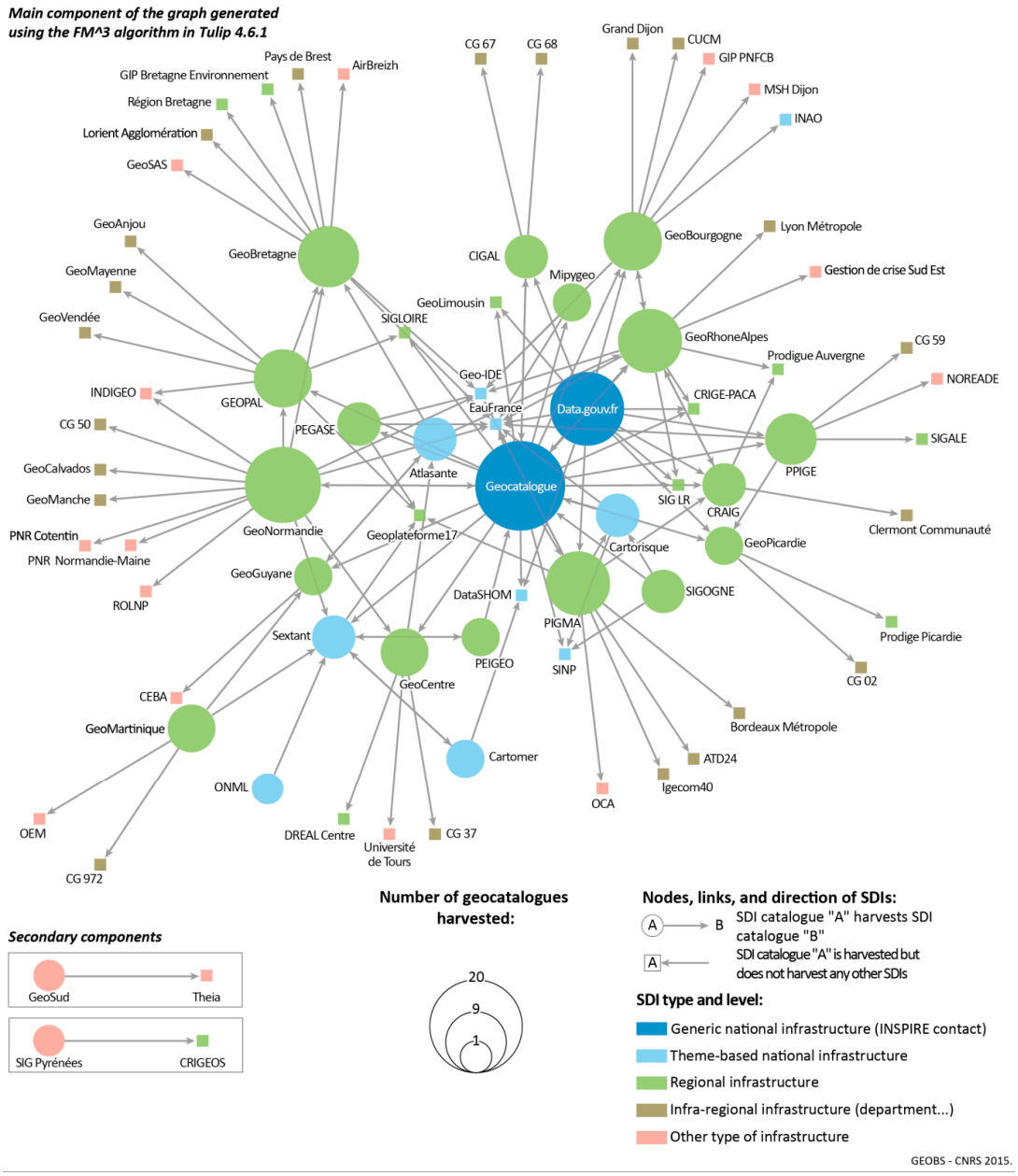


Figure 7. Network of extra-SDI links, established by analysing harvesting beyond the platforms included in this study.

The density of the graph (number of actual links/number of possible links) in Figure 7 is relatively low (d = 0.02). Regional SDIs have links to SDIs for departments or smaller administrative units, which have few connections to the rest of the network. The main component only includes 34 of the 45 SDI in the initial study group, which are associated with 45 other data portals (mainly on a local level). Capitalisation on the interoperability potential of these infrastructures is thus still relatively limited. The national geocatalogue, the INSPIRE focal point, is still the central node of the network. However, although the French government's open data portal is not, strictly speaking, an SDI, it is also strongly connected to the network.

3.3. Actor Networking

3.3.1. What Do the SDI Promoters Say about Geocollaboration Dynamics?

Most SDIs support the rollout of their technical system with a series of actions to promote the use of these tools and data by a broad spectrum of actors, and launch a dynamic around a particular topic or territory. Training in the use of tools and data, regulatory aspects, and specific themes are organised in parallel with communication actions and working groups. The aim is to launch active communities of practice to promote geocollaboration dynamics and involve a wide range of actors, even if there are significant differences between (more homogeneous) national and regional levels (Figure 8).

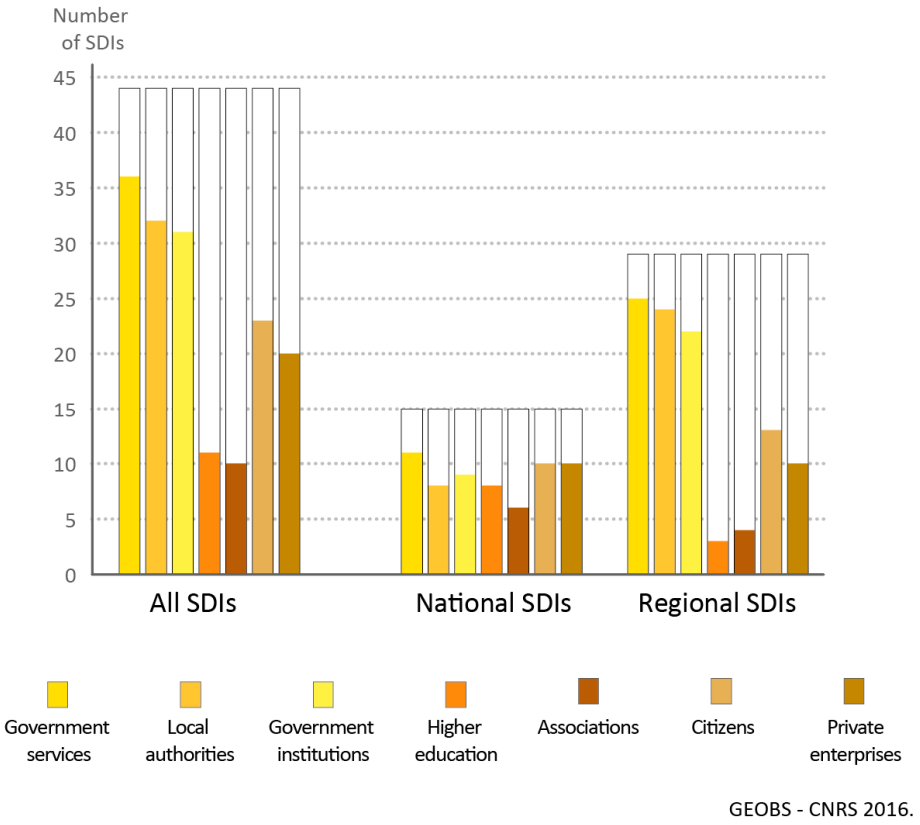


Figure 8. Distribution of the types of actors targeted by regional and national SDIs.

3.3.2. Do SDIs Promote Geocollaboration Dynamics?

Approximately half of the SDIs in the study organised communities of practice [20] to promote data coherence and sharing. While each SDI ran an average of 5–6 specialities, major differences were noted between national platforms (one third of which coordinated a speciality) and regional platforms (two thirds coordinated a speciality). The 132 functional specialities coordinated by the 45 SDIs in the study covered a very wide range of themes, giving an overview of the main current concerns of GIS specialists (Figure 9).

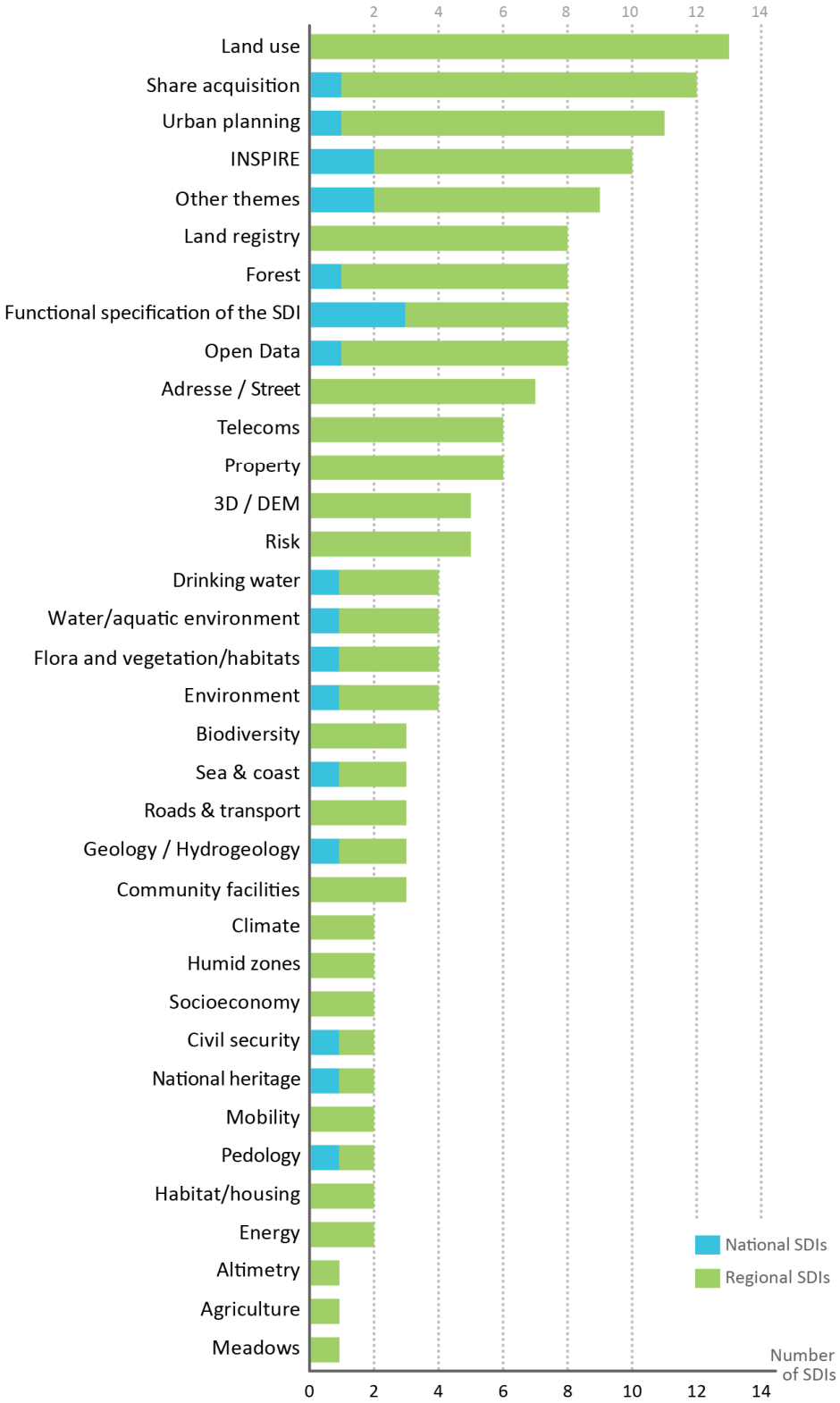


Figure 9. Distribution of working groups associated with national and regional SDIs in France.

3.3.3. Do SDIs Facilitate Networking among Diverse Actors?

Analysis of metadata fields concerning "contacts" (of all types) revealed little diversity among SDI contributors (Figure 10): Over 216,600 organisations were identified in the 160,000 SDI metadata

sheets studied (irrespective of their level) and 78% were public authorities (49% government services, 20% public bodies, and 9% local authorities). Teaching and research only represented 12% and associations and private businesses were in a very small minority (3%).

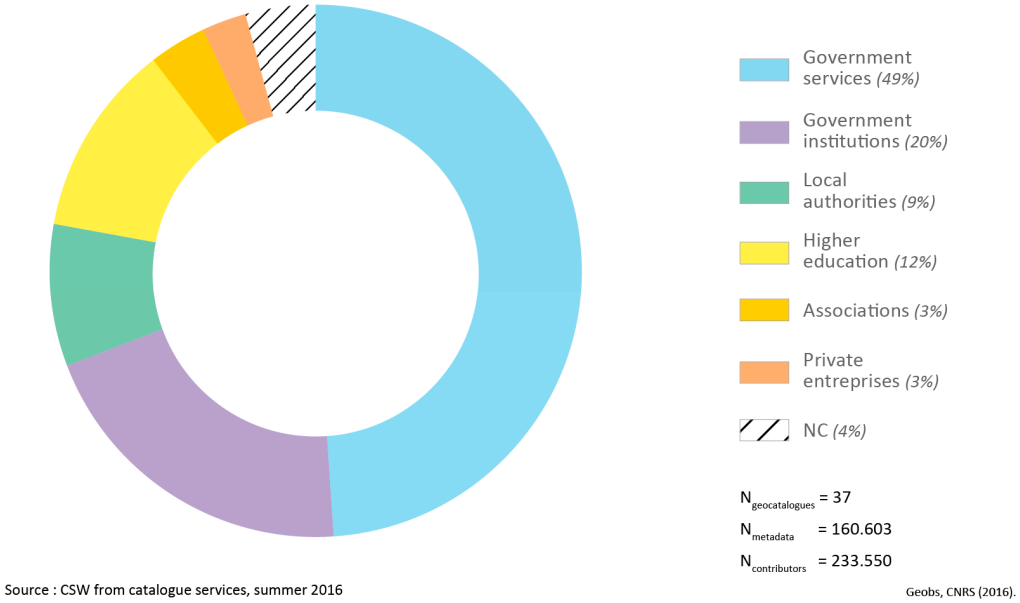


Figure 10. Frequency by type of organisation in French geocatalogue metadata.

The gap between levels was also perceptible, with a more marked presence of local authorities in regional SDIs and the scientific community in national SDIs.

3.4. Informational Equality of Territories

3.4.1. What Do SDI Promoters Say about Informational Equality of Territories?

An analysis of the spatial coverage of the sharable data in the national infrastructures under study, raises one of the recurrent arguments justifying the implementation of these SDIs: that of equal treatment of territories by governments in the form of homogenous coverage. This argument is often put forward to counterbalance the initiatives of multinationals, such as Google Maps, which only focus on areas with high commercial potential: "*The Géoportail covers the entire French territory, in accordance with the principle of equality and satisfying the general interest, while Google Maps covers the entire world with varying resolution according to the importance of the site*" (introduction page of the IGN's Géoportail, extracted from Wikipedia and translated by the authors).

3.4.2. Do SDIs Provide Homogeneous Coverage of Their Territories?

The extraction of bounding boxes from metadata makes it possible to question what a more "equitable" coverage of the territory would be like. Our map collections revealed the heterogeneity of information provided by the national and/or regional authorities, both for catalogue data (Figure 11) and catalogue and free-access data combined (Figure 12). Discretization by the equal interval method highlighted the informational inequalities between territories. Thus, the footprint of the oldest and most dynamic regional SDIs (Aquitaine (a), Pays de Loire (b)) and the most active local authorities (protected areas, such as the Bassin d'Arcachon (c)) are visible on general maps aggregating all the coverage. The latter thus demonstrates the importance of infra-national initiatives in the spatial distribution of the spatial data heritage, which has been indexed and, to a lesser extent, made available. In addition, the point of contact for the INSPIRE Directive clearly appears on these maps as the focal point of regional and national SDIs: the spatial patterns of regional and national

coverage are quite well aggregated on the map of Geocatalogue coverage. Finally, a comparison of Figures 11 and 12 also highlights the gaps between catalogue content in terms of metadata and really accessible spatial information.

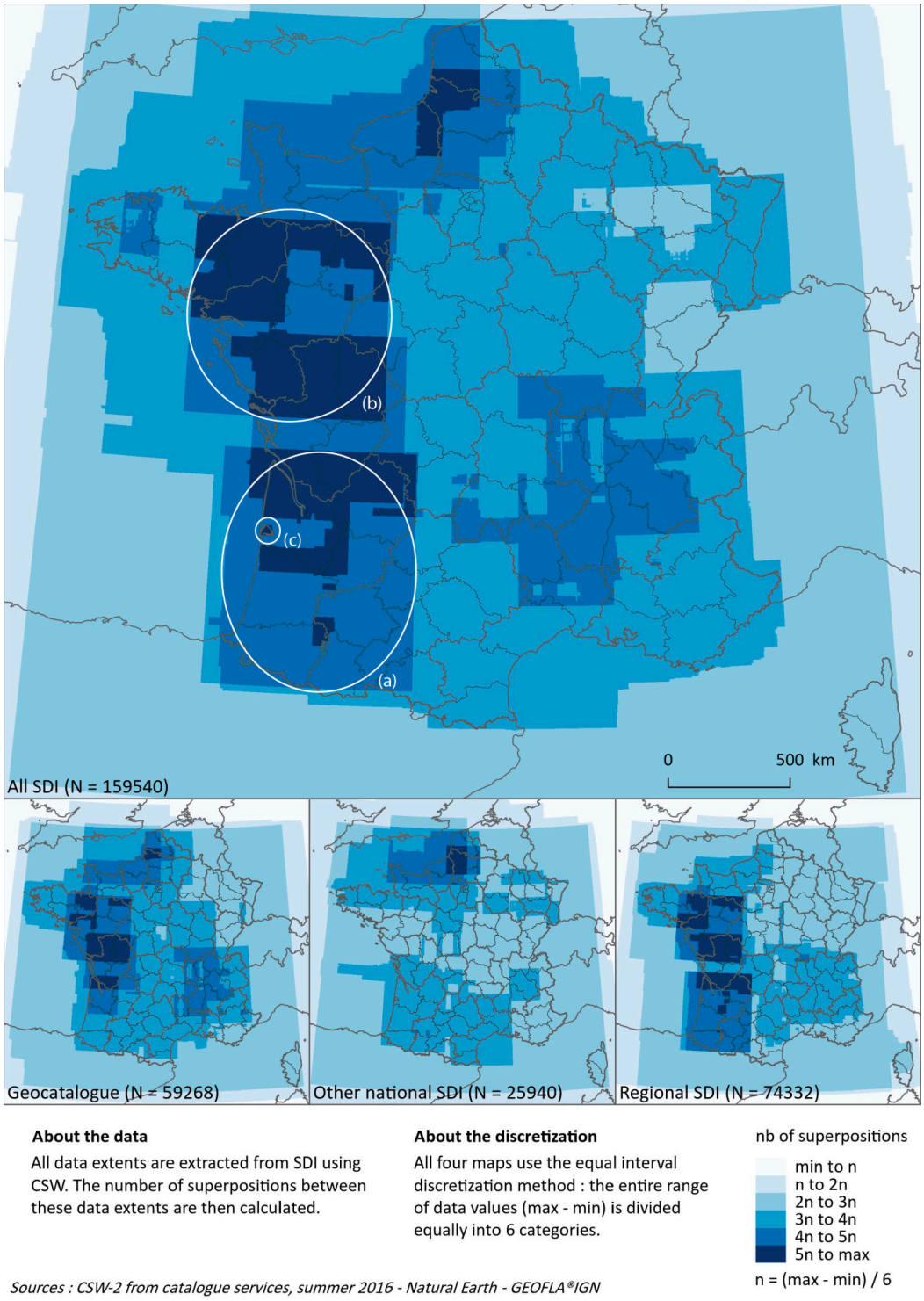


Figure 11. Mapping catalogue data coverage in French, national and regional SDIs.

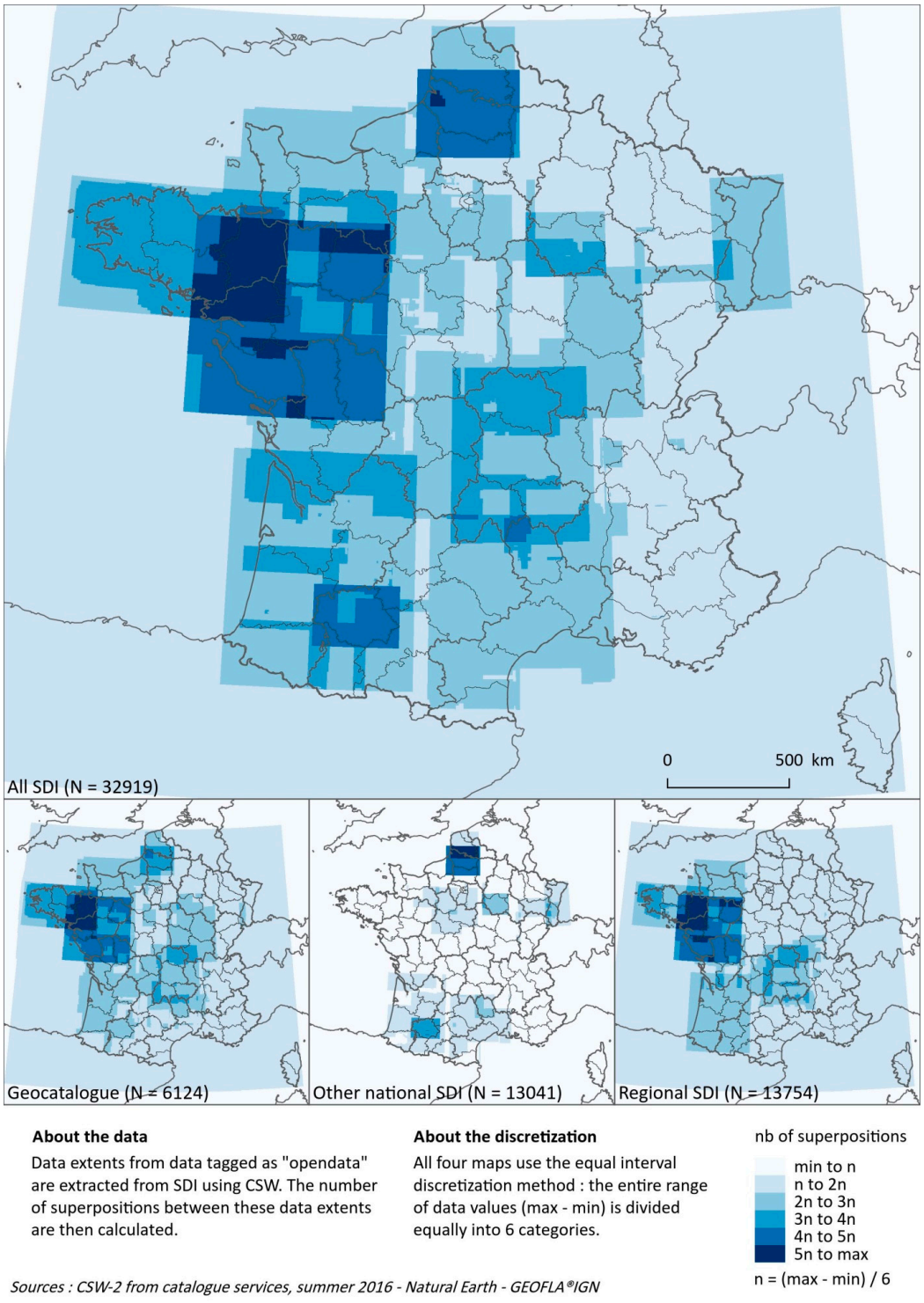


Figure 12. Mapping free-access catalogue data coverage in French, national and regional SDIs.

4. Discussion

This paper introduces a methodological framework for the study of the spatial data sharing through SDIs from two points of view, i.e., the discourse of their promoter and the reality of their data. The main motivation is to contribute to the critical analysis of this information technology [13]. Our study of French SDIs provides an overview of the four most important components of spatial data sharing.

While, globally, access to information is amplified for society, it becomes differentiated according to the territorial levels and the different actors sharing. As demonstrated by our survey coordinators, the lower is the level in the administrative hierarchy, the higher the restrictions in accessing a SDI become. This is partly due to the variation in the exchange forms between partners and the territorial scale. National platforms draw their legitimacy from their capacity to give access to a maximum of data to everyone, without users having to identify themselves (*open network*). A contrario, regional SDIs base their legitimacy on the fact that the actors of a network are clearly identified, and operate more easily using a closed network approach, limited to data exchange. In this regard, we conducted a survey on 45 French SDIs in the summer of 2015, and found that 60% of them required users to identify themselves and sign a charter before being able to access "advanced" functions. However, practices between national and regional levels are highly differentiated and even reversed: only one quarter of national platforms require signing a charter against three quarters of regional platforms. Based on a similar methodology (i.e., interviews with promoters of the SDIs using a questionnaire [21]) applied to almost the same list of French sub-national SDIs, the assessment conducted by Salgé et al. in 2009 [22] produced similar results: most of the SDIs analysed were only open to the public administration and only 6% were described as fully open to the general public. Six years later, there was no apparent trend towards greater openness. Confronted with the cost of organising and maintaining local digital infrastructures, SDIs which opened up during the 2000s are currently tempted, in France at least, to restore differentiated access favouring those who contribute to SDIs, or even those who are prepared to pay to access data. The open data movement is not, quite evidently, a process with no return, and the multiplication of digital "tolls" is also affecting the public field [5].

Furthermore, while interoperability has undoubtedly made it possible to achieve major progress in normalising web services and the potential interconnection of SDIs, our analyses of geocatalogue harvesting tend, however, to show that their actual scope is limited. Analysis of all the study material revealed that harvesting potential was still little used. On a European level, improving interoperability was one the five objectives of the INSPIRE Directive (2007/2/CE). However, a public consultation [23] carried out in 2014 by the European Commission found that this strand of INSPIRE was considered to be highly technically complex and required more support. Moreover, the mid-term evaluation report on INSPIRE implementation [24] highlighted the fact that "most of the measures to ensure interoperability have yet to be implemented", reflecting the French situation.

Although geocollaboration was studied from both design and technical perspectives, focusing on methods and tools to support the work of various groups based on spatial data (e.g., [25–27]), the stakeholders involved in these processes and the inter-organisational relationships had not received the same level of attention, thus failing to answer the question of how to initiate and maintain SDI networking activities among various stakeholders. Our results demonstrate that networking actors via collaboration dynamics while implementing SDIs seems to be effective. The 132 workgroups identified and the significant activity devoted to coordinating SDIs, particularly on a regional level, indicated that the communities of practice were very active. They played a key role in bringing together actors in a given territory to co-design models, co-produce data sets and/or co-control them. However, these geocollaboration dynamics are currently observed in an inward-looking system without diversification: government administrations and research act separately, the private sector is not represented, and associations are still only marginally involved.

Finally, regarding informational equality of territories, this research gives a whole new spatial perspective on the issue of the public access to environmental information [28]. An analysis of the coverage of reference data in metadata catalogues is useful for reviewing the public authorities' declarations on "egalitarian" treatment of the regions they administer. When many SDIs launched their platforms, they promoted the sharing of geographical frames of reference providing homogeneous coverage of their sphere of action. This logic is currently tending to be replaced by an approach aimed at aggregating the maximum amount of content, on a national, regional, or local level. Some striking contrasts have been revealed by this process, which may be indicative, not only of local issues, but also

of power struggles around administrative or regulatory boundaries. The difference between catalogue data and open data was also important, as it revealed major contrasts according to the area covered and the scale of analysis.

5. Conclusions and Future Prospects

The methodological framework proposed in this article compares the claims of SDI promoters (objectives and strategies) through interviews, a survey, and editorial analysis of the websites and their SDI content, based on an analysis of the metadata and the associated communities of practice. Assessment of the spatial and organisational coverage of the data, interconnections between SDIs, as well as the topics currently explored by communities of practice made it possible to explore four objectives highlighted by the sponsors of these projects: data accessibility, networking of tools, equal treatment of data in different territories, and geocollaboration dynamics. These four strategic objectives were divided into ten key questions, for a systematic analysis on a regional and national scale, in the case of France, of the gaps between the "promises of sharing" [5] spatial data and the actual implementation of SDIs. This analysis acknowledges the complex, dynamic, and multifaceted nature of SDIs. We implemented a multi-disciplinary research approach to understand this complexity and compare theory and practice. While SDI seems to be a positive "label" for government authorities, different forms of implementation are indicative of diverse technical and organisational realities. The methodological framework presented is the first stage in comparing declarations on SDIs with data actually shared and geocollaboration actions implemented. The second stage will consist of comparing the SDI promoters' assertions concerning the data and services and the actual use made of them [29]. As [7] quite rightly pointed out: "information should always be viewed within its context". We were aware of this but chose to focus in this article on the ways in which geographical information circulates and plan to focus on the determinants and effects of this circulation in a future project phase. The objective will be to assess the impact of spatial data sharing on regional management practices by a detailed analysis of actual uses and checking their coherence with the initial stated objectives of the SDIs studied. The aim of this contribution, therefore, is to provide a basis for analysing SDI content and promoters' claims, in order to understand, in the long term, how the various actors/SDIs compete for visibility of their spaces and/or topics on the Internet. Another perspective consists of studying the individual acceptance processes of SDIs, using, for example the Unified Theory of Acceptance and Use of Technology [30]. Finally, the entire project could be extended to cover a broader ecosystem, including SDIs, as well as all the open data portals and community projects (commercial or free-access) that are currently developing. The aim is to apply this methodological framework to analyse institutional SDIs to inverse infrastructure [8], i.e., user-driven (bottom up), self-organising infrastructure—resulting, for example, from science programmes involving the general public. The objective will be to determine whether the decentralised governance of spatial information introduces significant changes in data accessibility, interoperability of information systems, geocollaboration, and spatial equality.

Supplementary Materials: The histograms and maps presented in Figures 3, 4, 6, 7 and 9–12 will be available online on the research project website (http://geobs.cnrs.fr) when the article is published. It will thus be possible to explore the histograms interactively (choice of aggregated or detailed categories, displaying all the study material or the national and/or regional levels), as well as the graphs in Figures 6 and 7 (exploration of the neighbours of each node) and maps (choice of discretization and navigation) and the maps in Figures 11 and 12 (exploration of the spatial distribution of metadata). Once the article is published, this website will also provide access to the database, uploaded to http://data.gouv.fr under an Open Licence and the script, uploaded to http://github.com (see Appendixs A and B).

Acknowledgments: The authors are grateful to all the SDI coordinators for answering to the survey. This work was supported by the French *National Center for Scientific Research* (CNRS) and the Region *Nouvelle Aquitaine* under Grant 2014-1R40103.

Author Contributions: Matthieu Noucher and Françoise Gourmelon conceived the framework and wrote the paper; Adeline Maulpoix performed the survey; Julie Pierson and Mathias Rouan designed the metadata analysis; Olivier Pissoat designed the figures; Nathalie Pinède and Pierre Gautreau contributed to the strategic analysis; and Jade Georis-Creuseveau discussed the results.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. CSW-Harvester

Appendix A.1. Synopsis

This is a python script that harvests metadata from CSW web services and saves some information from these metadata in a postgreSQL database.

Appendix A.2. Motivation

This script is used to analyse Spatial Data Infrastructures for the GEOBS research project: https://www-iuem.univ-brest.fr/pops/projects/geobs.

Appendix A.3. Dependencies

- psycopg2: PostgreSQL database adapter for Python, https://pypi.python.org/pypi/psycopg2
- OWSlib: Python package for client programming with Open Geospatial Consortium (OGC) web service (hence OWS) interface standards, https://github.com/geopython/OWSLib
- OWSlib requires elementtree (standard in 2.5 as xml.etree) or lxml
- PostgreSQL

Appendix A.4. How to Run

The PostgreSQL database must first be created. A database dump is provided with database/csw_harvester.sql.

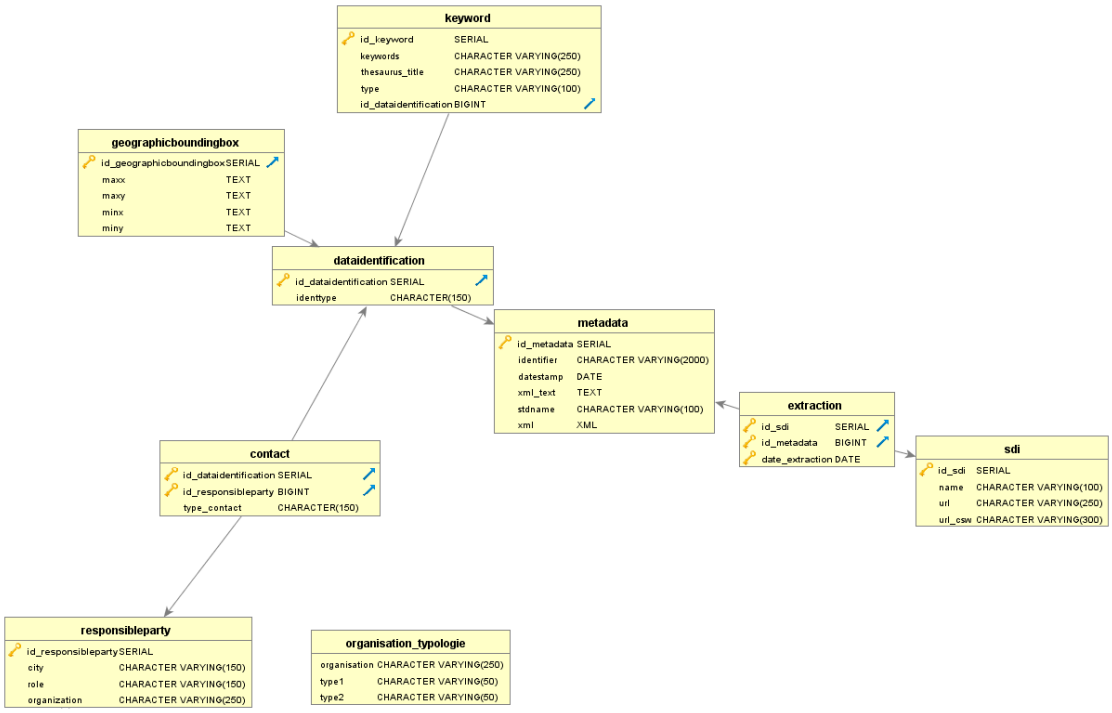


Figure A1. Conceptual Model.

The host, port, database name, schema, user and password must be set in csw-harvester.py. The CSW list is read from a CSV file; an example is provided with sources-csw.csv. For each CSW, you can set a start and end record, and the number of records extracted in each step (for example, if set at 30, records will be extracted 30 by 30). Lines can be commented with #.

You can then run the python script csw-harvester.py with the following options:

- -f OUTPUTSCHEMA, –outputschema=OUTPUTSCHEMA outputschema: default = http://www.isotc211.org/2005/gmd
- -s SOURCES, –sources=SOURCES CSV file, 7 fields separated by a comma: num_idg, name_idg, begin_record, end_record, MAXR, url, url_idg
- -l LOG_FILE, –log-file=LOG_FILE LOG file
- -c, –completion completion mode
- -d DATE, –date=DATE Extraction date

The completion mode (true by default) is used to run another iteration of the script without overwriting the data already stored in the database. This is useful if, for example, one CSW stopped working after a given record and you want to start again from this record.

The date option is used to force the extraction date stored in the database, if, for example, you are using the completion mode and want the extraction date to be the same for all metadata.

Appendix A.5. License

This project is published under the General Public License v3.

Appendix A.6. URL

https://github.com/LETG/csw-harvester.

Appendix B. Medata-Extents

Appendix B.1. Synopsis

This is a python script that creates an ASC raster layer counting superpositions of rectangular extents defined by four coordinates.

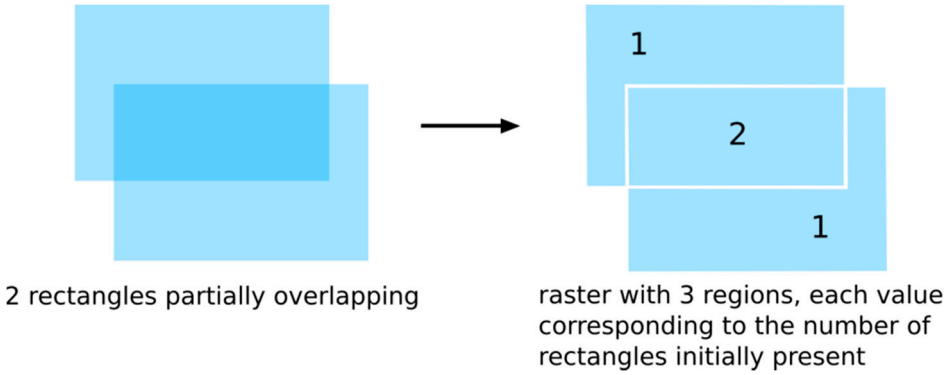


Figure B1. General Principle.

The resulting ASC layer can be viewed in GIS software to produce this kind of map:

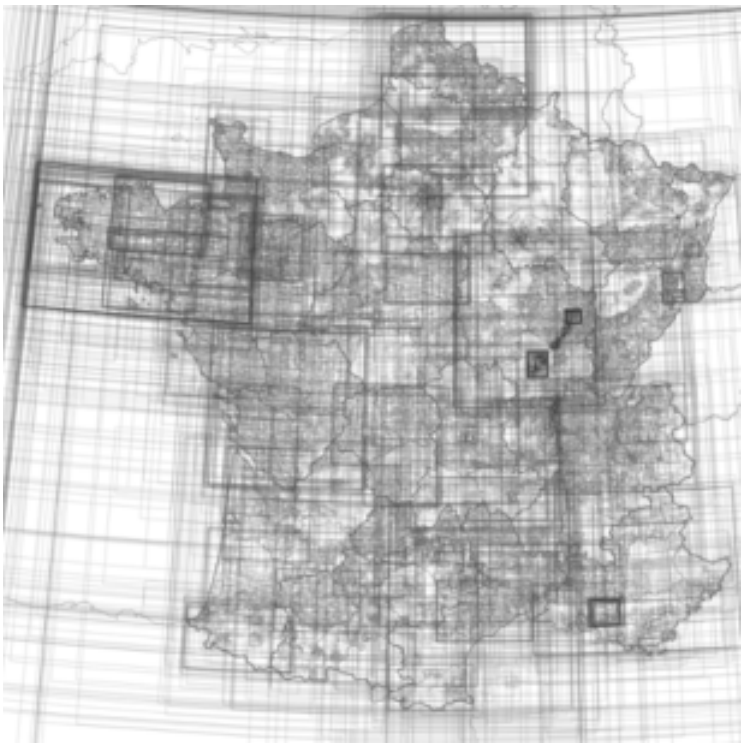


Figure B2. Example of metadata extents.

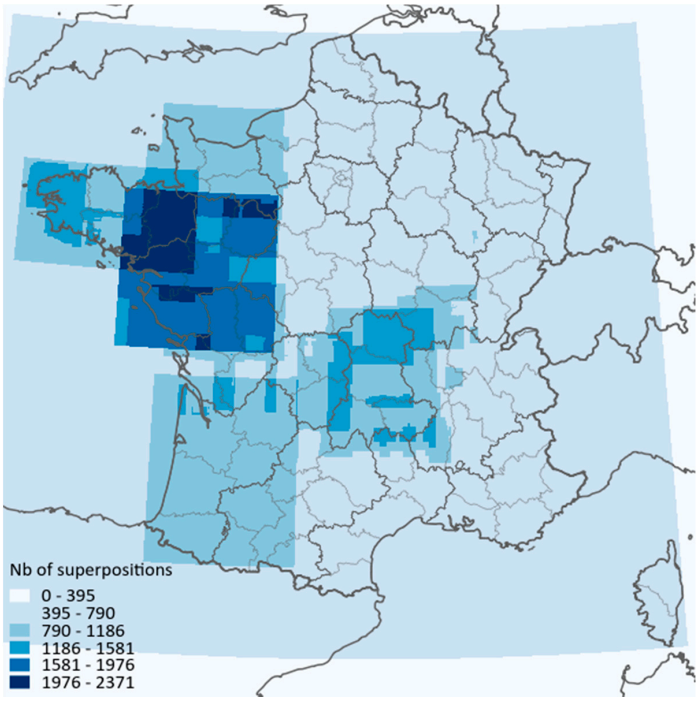


Figure B3. Example of a map produce by the script.

Appendix B.2. Motivation

This script is used to analyse Spatial Data Infrastructures for the GEOBS research project: https://www-iuem.univ-brest.fr/pops/projects/geobs

Appendix B.3. Dependencies

- numpy: package for scientific computing with Python, http://www.numpy.org/
- pandas: Python data analysis library, http://pandas.pydata.org/

Appendix B.4. How to Run

This script uses a CSV file as an input, with each line corresponding to one rectangular extent, and four columns for the four coordinates of each extent (west, east, north, south). These four columns can be in any order and labelled as you wish. Open metadata_extents.py with a text editor and set the variables at the beginning of the script:

- resolution (same unit as coordinates used in CSV, i.e., degrees if coordinates in WGS84)
- path to CSV
- delimiter of CSV
- quotechar of CSV
- names of CSV columns
- name of output raster to be created
- name of log file

Once you have set these variables, run the script (http://pythoncentral.io/execute-python-script-file-shell/).

Appendix B.5. License

This project is published under the General Public License v3.

Appendix B.6. URL

https://github.com/UMR-PASSAGES/metadata-extents.

References

1. Craglia, M.; de Bie, K.; Jackson, D.; Pesaresi, M.; Remetey-Fülöpp, G.; Wang, C.; Annoni, A.; Bian, L.; Campbell, F.; Ehlers, M.; et al. Digital Earth 2020: Towards the vision for the next decade. *Int. J. Dig. Earth* **2012**, *5*, 4–21. [CrossRef]
2. Rajabifard, A.; Feeney, M.-E.; Williamson, I.; Masser, I. Chapter 6, National SDI Initiatives. In *Development of Spatial Data Infrastructures: From Concept to Reality*; Williamson, I., Rajabifard, A., Feeney, M.-E.F., Eds.; Taylor & Francis: London, UK, 2003; pp. 95–109.
3. Crompvoets, J.; Bregt, A.; Rajabifard, A.; Williamson, I. Assessing the world wide developments of national spatial data clearinghouses. *Int. J. Geogr. Inf. Sci.* **2004**, *18*, 665–689. [CrossRef]
4. Masser, I. *Building European Spatial Data Infrastructures*; ESRI Press: Redlands, CA, USA, 2010.
5. Gautreau, P.; Noucher, M. Sharing Platform in Digital Geographic Information: Everything It Promise? *Justice Spat./Spat. Justice* **2016**, *10*. Available online: http://www.jssj.org/ (accessed on 27 March 2017).
6. Mol, A.P.J. Environmental governance through information: China and Vietnam. *Singap. J. Trop. Geogr.* **2009**, *30*, 114–129. [CrossRef]
7. Kooper, M.N.; Maes, R.; Lindgreen, E.R. On the governance of information: Introducing a new concept of governance to support the management of information. *Int. J. Inf. Manag.* **2011**, *31*, 195–200. [CrossRef]
8. Mol, A. *Environmental Reform in the Information Age. The Contours of Informational Governance*; Cambridge University Press: Cambridge, UK, 2008.
9. Hendriks, P.H.J.; Dessers, E.; Van Hootegen, G. Reconsidering the definition of spatial data infrastructure. *Int. J. Geogr. Inf. Sci.* **2012**, *26*, 1479–1494. [CrossRef]
10. Coetzee, S.; Wolff-Piggott, B. A Review of SDI Literature: Searching for Signs of Inverse Infrastructures. Cartography—Maps Connecting the World. In *Part of the series Lecture Notes in Geoinformation and Cartography*; Springer International Publishing: Zurich, Switzerland, 2015; pp. 113–127.

11. Percivall, G. The application of open standards to enhance the interoperability of geoscience information. *Int. J. Dig. Earth* **2010**, *3*, 14–30. [CrossRef]
12. Genovese, E.; Cotteret, G.; Roche, S.; Caron, C. Evaluating the socio-economic impact of geographic information: A classification of the literature. *Int. J. Spat. Data Infrastruct. Res.* **2009**, *4*, 218–238.
13. Reznik, T.; Chudy, R.; Micietova, E. Normalized evaluation of the performance capacity and availability of catalogue services: A pilot study based on Infrastructure for Spatial Information in Europe. *Int. J. Dig. Earth* **2016**, *9*, 325–341. [CrossRef]
14. Ballatore, A. Exploring the geographic information universe: The role of search technologies. In Proceedings of the Workshop on Geographic Information Observatories at GIScience, Vienna, Austria, 23–26 September 2014.
15. Kitchin, R.; Lauriault, T. Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work. In *Geoweb and Big Data*; Eckert, J., Shears, A., Thatcher, J., Eds.; University of Nebraska Press: Lincoln, NE, USA; Available online: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2474112 (accessed on 27 March 2017).
16. Ladurelle-Tikry, E.; Debord, M. French regional SDI Network: Parternship for sharing Inspire experience. In Proceedings of the INSPIRE Conference, Istanbul, Turkey, 23–27 June 2012.
17. Noucher, M.; Gourmelon, F.; Georis-Creuseveau, J. Toward an Observatory Prototype of the Contents of the 65 french SDIs. In Proceedings of the Inspire Conference—Geospatial World Forum, Lisbon, Portugal, 25–29 May 2015.
18. Johnson, R.B.; Onwuegbuzie, A.J. Mixed Methods Research: A Research Paradigm whose Time Has Come. *Educ. Res.* **2004**, *33*, 14–26. [CrossRef]
19. Devillers, R.; Jeansoulin, R. *Fundamentals of Spatial Data Quality*; ISTE Publishing Company: Eugene, OR, USA, 2006; p. 312.
20. Wenger, E. *Communities of Practice: Learning, Meaning, and Identity*; Cambridge University Press: Cambridge, UK, 1998.
21. Rix, J.; Fast, S.; Masser, I.; Salgé, F.; Vico, F. Methodology to Describe, Analyse and Assess Subnational SDIs: Survey, Experiences and Lessons Learnt. *Int. J. Spat. Data Infrastruct. Res.* **2011**, *6*, 23–52.
22. Salgé, F.; Ladurelle-Tikry, E.; Fourcin, L.; Dewynter, B. Review of sub-National SDIs in France An outcome of the eSDI-Net+ project. In Proceedings of the GSDI Conference, Rotterdam, The Netherlands, 15–19 June 2009.
23. Craglia, M.; Roglia, E.; Tomas, R. INSPIRE Public Consultation 2014, Reports on Findings. Technical Report by the Joint Research Centre of the European Commission. 2014. Available online: http://inspire.ec.europa.eu/reports/consultations/INSPIRE_Public_Consultation_Report_final.pdf (accessed on 27 March 2017).
24. European Environment Agency. Mid-Term Evaluation Report on INSPIRE Implementation. EEA Technical report No 17/2014. 2014. Available online: http://www.eea.europa.eu/publications/midterm-evaluation-report-on-inspire-implementation (accessed on 27 March 2017).
25. Balram, S.; Dragicevic, S. Geocollaboration. In *Encyclopedia of GIS*; Shekhar, S., Xiong, H., Eds.; Springer: New York, NY, USA, 2008; pp. 344–347.
26. Chang, Z.; Li, S. Geo-social model a conceptual framework for real-time geocollaboration. *Trans. GIS* **2013**, *17*, 182–205. [CrossRef]
27. Sun, Y.; Li, S. Real-time collaborative GIS: A technological review. *ISPRS J. Photogramm. Remote Sens.* **2016**, *115*, 143–152. [CrossRef]
28. United Nations Economic Commission for Europe, Convention on Access to Information, Public Participation in Decision-Making and Access to Justice in Environmental Matters. 1998. Available online: http://www.unece.org/fileadmin/DAM/env/pp/documents/cep43e.pdf (accessed on 27 March 2017).
29. Georis-Creuseveau, J.; Claramunt, C.; Gourmelon, F. A modelling framework for the study of Spatial Data Infrastructures applied to coastal management and planning. *Int. J. Geogr. Inf. Sci.* **2017**, *37*, 122–138. [CrossRef]
30. Venkatesh, V.; Morris, M.G.; Davis, G.B.; David, F.D. User acceptance of information technology: Toward a unified view. *MIS Q.* **2003**, *27*, 425–478.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).