An Evaluation of Semantically Enriched Spatial Data Infrastructure

Hamon Barros Henriques
Computer and Systems Depart.
Federal Univ. of Campina Grande
Campina Grande, Brazil
hamon@copin.ufcg.edu.br

Fabio Gomes de Andrade
Computer and Systems Depart.
Federal Institute of Paraíba
Cajazeiras, Brazil
fabio@ifpb.edu.br

Cláudio de Souza Baptista
Computer and Systems Depart.
Federal Univ. of Campina Grande
Campina Grande, Brazil
baptista@dsc.ufcg.edu.br

***Abstract*— Spatial Data Infrastructures (SDI) have become an important solution for easing the interoperability of geographic data offered by different organizations. An important challenge that must be overcome by such infrastructures consists in allowing their users to easily locating the available data. Presently, this task is implemented by means of catalog services, which still have important limitations that prevent effective data retrieval. Due to those limitations, many research works have been developed to improve information retrieval in SDIs. One of such works is Semantically Enabled Spatial Data Infrastructures (SESDI), which is a framework that uses a model-based on information at feature type level and ontologies. The first results obtained during the experimental evaluation of SESDI showed that it improved the quality of several kinds of queries concerning geographic data. Nevertheless, a deeper evaluation, besides the comparison to catalog services provided by other infrastructures, was still necessary. Aiming at meeting this need, this paper describes an experiment carried out in order to deepen that evaluation. In this experiment, the performance of SESDI was compared with catalog services offered by other two infrastructures. The results obtained from the new experiments showed the viability of the solution used to implement the framework.**

Keywords-spatial data infrastructure; catalog service; spatial database; GIS.

I. INTRODUCTION

Recently, Spatial Data Infrastructures (SDI) have become popular as an important solution for allowing the interoperability of geographic data supplied by different organizations. In order to achieve this interoperability, these infrastructures are created based on a set of norms and standards that must be adopted by all of their components.

In this standardization scenario, the standards specified by the Open Geospatial Consortium (OGC) [1] have played a key role. Important examples of these standards include the services that give support for accessing geographic data, such as Web Map Service (WMS) [2] and Web Feature Service (WFS) [3]. These services allow the clients to have access to geographic data offered by a given provider using a standard interface, with no need to be aware of the details about its storage. Each instance of these services gives access to a set of feature types. On the other hand, each feature type represents a layer present in the dataset offered by a provider. Since their proposal, the OGC web services have played a key role for the development of spatial data structures, being used in the implementation of many of the current infrastructures.

Besides interoperability, an important challenge which must be overcome by SDIs consists in allowing the clients to easily locating the available data and services. Presently, this problem is usually solved by the implementation of catalog services. In a catalog service, geographic data providers must supply a metadata containing detailed information about their dataset. On the other hand, clients of a SDI may use this service in order to find out the geographic data they are interested in.

Although having led to important advances, the present catalog services still have some limitations. One of these limitations is that they solve their queries solely based on the information contained in the metadata records created by the geographic data providers. Since these records normally describe the datasets offered by a service as a whole, they usually bring little or even no information concerning the spatial, thematic and temporal characteristics of each feature type, which constrains the information retrieval process. Another important limitation is that most of the present catalogs solve queries with thematic constraints based on keywords only. This characteristic causes the catalog to discard, during the information retrieval process, relevant resources that are described with terms related to the theme defined in the query, which reduces the recall of these queries.

Throughout the years, many research works have been proposed with the objective of overcoming these limitations. One such work is Semantically Enabled Spatial Data Infrastructures (SESDI) [4]. SESDI is a framework that uses a model-based on metadata at feature type level and ontologies to improve the retrieval of geographic data offered by an SDI. During the evaluation of this framework, the results achieved by its search engine were compared with the results achieved by similar queries submitted to the catalog service offered by a present SDI. This validation, which used the North-American SDI (NSDI) [5] as case study, showed that SESDI improved the quality of several kinds of queries. Nevertheless, a deeper evaluation, besides the comparison to other available catalog services, was still necessary.

Aiming at meeting this need, this paper describes an experiment carried out to deepen the evaluation of SESDI. In this experiment, the performance of SESDI was compared with the catalog services offered by other two infrastructures, namely the Canadian Spatial Data Infrastructure [6] and the Global Earth Observation System of Systems (GEOSS) [7]. During the experiment, each solution was evaluated based on four query types: spatial, thematic, temporal and global. Besides allowing a deeper evaluation of SESDI, this paper offers as contribution an analysis of the performance of the studied catalog services, evaluating their quality at solving several kinds of queries involving spatial data.

The remaining of this paper is organized as follows. In Section II, we approach related works. Section III gives an overview of the two approaches studied in this work. Section IV focuses on the experiment design, describing the dependent variables, the response variables, the experimental units, the research hypotheses and the data collection process. Section V describes the results achieved with the experiment. Finally, Section VI concludes the paper, and discusses further work to be undertaken.

II. RELATED WORKS

Recently, many works have been proposed with the objective of improving the retrieval of geographic data and services from spatial data infrastructures and geographic portals. Some of these works are concerned with the retrieval of services. Stock et al. [8] developed a solution in which the services are retrieved from information defined in a feature type catalog that defines the structure and the relationships among the feature types offered by the SDI. Lutz et al. [9] associated the output parameters of each service to concepts defined in application ontologies to improve information retrieval in disaster management applications. In another work, Lutz [10] used first order logic to generate a signature which describes the semantic of the services offered by an SDI. In the approach developed by Lemmens et al. [11], services were classified and retrieved according to a geospatial web services ontology. Other proposed solutions [12][13] described the available services as tasks. Li et al. [14] proposed a solution which expands the terms defined in the query of a client in order to improve the retrieval of geographic web services that offer data about the Arctic region. The disadvantage of all these works is that they do not take into account the information offered at feature type level during the resolution of queries.

Some of the developed works manage to solve queries based on information at feature type level. Janowicz et al. [15] developed a solution which used a semantic similarity measure based on ontologies to retrieve geographic data. In another work, Zhang et al. [16] proposed a solution that resolves queries for feature types using specific ontologies related to four types of dimension: location, theme, geometry and properties. Li et al. [17], in turn, implemented a solution that resolves thematic queries at feature type level using a keyword-based approach, which constrains the quality of this kind of query. A disadvantage of all these studies is that they do not approach the resolution of temporal queries, which are important for many queries involving geographic data.

Finally, some proposed works deal with the retrieval of geographic metadata, and do not focus on a specific kind of resource. Smits and Friis-Christensen [18] developed a solution which uses a multilingual thesaurus to describe the information offered in the catalog of an infrastructure. Athanasis et al. [19], in turn, developed a solution that describes the resources offered by an SDI by means of metadata based on a series of ontologies. Chen et al. [20] used OWL-S to describe the semantics of data offered through geographic web services. Despite their importance and relevance, these works cannot solve queries based on information at feature type level. Also, these solutions do not support the resolution of temporal queries.

III. CATALOG SEVICES VERSUS SESDI

This section provides an overview of the two approaches that were compared during the experiment. First, we show how the queries are solved by the current catalog services. After that, we show the approach used in the implementation of SESDI.

A. Catalog Service

Currently, geographic data providers use the catalog service to advertise their resources. For this, they create a new metadata record into the catalog service. During this process, they supply metadata that describe their datasets, according to the geographic metadata standard adopted by the SDI. Some examples of these metadata include the title of the dataset, the coordinate systems and the projection used in the production of data, and the URL from which the service can be invoked. The metadata provided by each metadata record are used by the catalog service during the query resolution process. Since geographic data providers normally use a single metadata record to describe their datasets as a whole, current catalog services have limitations to solve spatial, thematic, temporal and global queries.

Regarding spatial queries, the main limitation is that data providers normally define a single bounding-box to represent their datasets. Then, during the searching process, the catalog selects all the records whose bounding-box intersects (or contains) the geographic region defined in the query. Nevertheless, it is possible to notice that sometimes the feature types provided in a dataset cover different regions, which causes some limitations to the searching process. To understand these limitations, suppose the case depicted in Figure 1, which describes a dataset about flooding. In that figure, the feature types present in the dataset are represented as green rectangles, while the blue rectangle represents the spatial, thematic and temporal information provided in the metadata record that describes the dataset. Also, suppose B1 and B2 are two bounding-boxes covering two different regions that do not overlap each other. In Figure 1, the geographic extent provided in the record is B1, which represents the region that is covered by most of the feature types of the dataset. Then, if a user poses a query looking for maps about the region B2, the catalog service discards the record depicted in that figure, even its dataset has a feature type that satisfies the search criteria. Moreover, if a user poses a query for maps about the region B1, the catalog returns the record. After that, the user needs to access the service to identify, among the entire feature types of the dataset, just the ones that are relevant to his/her query. This process can be quite tedious and tiring for the user, since a query may return a large volume of records and each service, in turn, may offer a large number of feature types. The limitations regarding spatial queries can be extended to temporal queries. Moreover, in temporal queries these limitations are even bigger, since many providers do not provided the metadata about the temporal extent of the dataset.

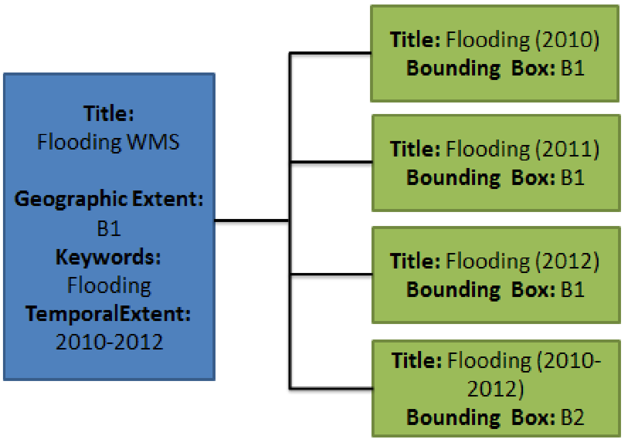

Figure 1. Example of a metadata record.

Regarding thematic queries, the main limitation is that catalog normally solves queries based on keywords. During the searching process, the catalog selects the records that contain in their description the keyword used in the query. Then, relevant records that contain in their description keywords that are related to the theme requested in the query can be discarded during this process. For example, suppose the case depicted in Figure 1 again. If a user poses a query for maps about *disasters*, the record showed in the figure is not selected by the catalog service, even its dataset offers several feature types that are relevant to the user. Finally, global queries are even more difficult to solve by using the current catalogs, since they present the limitations concerning all the dimensions used in the query.

B. SESDI

SESDI is a framework proposed with the objective of easing the retrieval of geographic data offered by SDIs, and aims to overcome some of the limitations of the present catalog services. In order to achieve this goal, it solves the queries of the clients based on a model that adapts classic Information Retrieval (IR) techniques to the domain of geographic data. In this model, the services that offer access to geographic data, such as WMS and WFS, are described as a set of feature types, in the same way as the documents are described as a set of keywords in the classic information retrieval models. Another important characteristic of SESDI is that it uses ontologies to describe the semantic of the feature types offered by the service, in order to improve the quality of the queries with thematic constraints. Finally, the framework proposes a search engine that explores the spatial, thematic and temporal relationships among the feature types offered by a service in order to generate the results retrieved in their queries.

In order to implement its model, the SESDI identifies the spatial, thematic and temporal characteristics of each feature type present in the dataset offered by the services registered in the catalog service of the infrastructure. To perform this task, it processes the information of each metadata record in order to retrieve information concerning the services offered by the SDI. After that, it invokes the *GetCapabilities* operation of each identified service in order to retrieve its capabilities document. The objective of this stage is to retrieve more detailed information about its feature types.

At the end of these stages, the framework processes the information contained in the metadata record and in the capabilities document to identify the spatial, thematic and temporal characteristics of each feature type. This entire process is called tagging and is split into three stages: spatial tagging, thematic tagging and temporal tagging. It is important to take in mind that the tagging process is executed for each feature type identified by the framework. The following paragraphs describe the tagging process. However, more detailed information about this process can be found in [4].

In the spatial tagging, the SESDI identifies the geographic region covered by the feature type. That information, which is represented as a bounding box, is obtained from the feature type description in the capabilities document retrieved from the service.

In the thematic tagging, the framework tries to identify the semantics of the data provided by the feature type. It accomplishes that task by relating the feature types to concepts defined in ontologies. In the current version, the SESDI uses a set of ontologies about several application domains. During the thematic tagging, the framework processes the keywords list provided for the feature type in the capabilities document of the service. If no keywords are provided for the feature type, its title is used as the input for that process. Then, the SESDI matches each keyword (or the title of the feature type) to the names of the concepts used in its ontologies. Whenever a match is identified, the framework generates a tag associating the matched concept to the feature type. When a matching cannot be identified, the SESDI poses a query in the Wikipedia and tries to retrieve a page related to the keyword. If more than one page is retrieved, a ranking for each page is generated by using techniques of classic information retrieval, and the page with the highest ranking value is selected. After that, the framework matches the title of the selected page, as well the name of each category used in the page description, to the concepts used in its ontologies. The thematic tags generated during this process are associated to the feature type description and stored in a database. If no matches are identified during the tagging process, the framework does not generate any tag to the feature type.

Finally, the temporal tagging consists of identifying the time period of the feature type. Since the capabilities document does not provided metadata specific to describe temporal information, the SESDI processes the title and the text description of the feature type in order to find temporal expressions that provides that information. When one or more temporal expressions are found, their values are converted into a time interval. On the other hand, if the framework cannot identify any temporal expression, it assumes that the time interval of the feature type is the same defined in the metadata record. Furthermore, if no value is provided for temporal extent in the metadata record, the value null is used as the time interval of the feature type. After the tagging process, the metadata generated by SESDI are stored in a local database, along with some description about the feature type and its respective service. Figure 2 shows the metadata generated by SESDI for the feature types depicted in Figure 1 after the tagging process.

A key difference between SESDI and the present catalog services is that SESDI solves queries based on the metadata identified for each feature type during the tagging process. Another important difference is that SESDI returns to the users only the feature types that satisfy the search criteria, so clients do not need to access the service to identify the feature types of his/her interest. For example, if a user poses a spatial query for maps about the region B2, the framework is able to identify that the feature type "*Flooding (2010-2012)*" is relevant for the query. Moreover, if a user poses a query for maps about the region B1, the framework returns only the feature types that satisfy the constraint defined in the query. Another important difference is that the SESDI uses thematic tags associated to concepts defined the ontologies. Then, when a user poses a query for a specific theme, the framework is able to return as the feature types that are tagged exactly with the theme used in the query as the ones that are tagged with a concept related to it.

So, the study described in this paper was intended to carry out a deeper evaluation of the performance of these two kinds of approach. This approach was based on an empirical experiment, where we defined the controlled variables, which are the variables subject to adjustments before the execution of the experiment, and the dependent variables, which result from the experiment. Besides, with these variables, we elaborated the hypotheses to be verified. The next sections describe, respectively, the experiment design and the evaluation of the results.

IV. Experiment Design

The experiment followed the same strategy of the evaluation used in [4], which employed recall and precision metrics to evaluate the performance of SESDI with respect to catalog services of the Canadian SDI and of the GEOSS. For each kind of query (spatial, temporal, semantic and global), the performance was measured according to the recall (number of relevant results) and precision (quality of the relevance) of the results. On the other hand, at the service level, recall and precision were evaluated with basis on the number of retrieved services.

Figure 2. Example of feature type metadata generate by SESDI.

The kinds of queries used in the research are classified as purely spatial, purely semantic, purely temporal and global. The purely spatial queries are intended to search for maps that intersect, partially or entirely, a certain location of interest to the user (e.g., which cities are crossed by the São Francisco river?). The purely semantic queries, in turn, are intended to retrieve maps that describe geographic data about a certain theme (subject) (e.g., which maps refer to beaches?). The purely temporal queries have the objective of retrieving the maps that refer to a certain time interval (e.g., *which maps refer to the decade of 1950*?). Finally, the global queries retrieve maps that deal with a certain theme in a certain location and in a time interval, that is, these queries meet more than one constraint type (e.g., *which maps refer to beaches in Brazil in the decade of 1950*?).

The controlled variables used in the experiment were: kind of query used (purely spatial, purely temporal, purely semantic or global) and the used tools (SESDI and the catalog services of the Canadian SDI and of the GEOSS). On the other hand, the dependent variables used to compare the performance of SESDI with respect to the studied SDIs were recall and precision. These are the main metrics for evaluation of the performance of the information retrieval system. The recall is measured as the ratio between the number of relevant resources retrieved and the number of relevant resources that exist in the system. The precision is measured as the ratio between the number of relevant resources retrieved and the total number of resources retrieved.

The experimental units were composed by the comparisons of SESDI with the Canadian SDI and the GEOSS. For each of them, twenty queries of each type were performed [21]. For each round of testing, the query that was held in the SESDI was held too in the SDIs studied. Moreover, during the comparison between the SESDI and the SDIs, the queries were held on the same set of metadata.

Before the formulation of the hypotheses, the concepts of independent variables and response variables were formalized, in order to provide a better visualization of the presented information. The independent variable "query type" was formalized by the symbols *e*, *s*, *t* and *g*, for the spatial, semantic, temporal and global levels, respectively, while the independent variable "used tool" was formalized by the symbols *sd*, *c* and *g*, representing SESDI, the Canadian SDI and the GEOSS SDI, respectively. In order to represent the response variables "precision" and "recall", the following functions were formalized:

- *PFT*: represents the precision at feature type level, which evaluates the quality of the feature types retrieved from each approach;
- *PS*: represents the precision at service level, which evaluates the quality of the services retrieved from each approach;
- *CF*: represents the recall at feature type level, which evaluates the quantity of relevant feature types retrieved from each approach; and
- *CS*: represents the recall at service level, which evaluates quantity of relevant services retrieved from each approach.

The objective of this study was to evaluate the performance of SESDI with respect to the catalog services provided by some existing SDIs. The null hypotheses formalized to describe the comparisons between SESDI and the SDIs used as case studies are depicted in Table I. The first column describes if the hypothesis represent a query at level of service or feature type. The second and third columns represents, respectively, the null hypotheses used during the comparison between SESDI and the catalog service provided by the Canadian SDI and GEOSS. In Table I, each hypothesis represents a comparison between the performance of SESDI and the catalog service provided by a SDI. These hypotheses were formulated based on the metrics presented above (*PFT*, *PS*, *CF* and *CS*). For example, the hypothesis $H_0 1$assumes that the precision at feature type level in the spatial queries held in SESDI is less or equal than the precision at feature types level in the spatial queries held in Canadian SDI. For each hypothesis defined in Table I, we performed statistical tests to try to refute it. The results of statistical tests are showed in Tables II and III.

The process of collecting the research data occurred in a semi-automatic and individual manner for each one of the queries.

The first step was the generation of the baseline. In this stage, the relevant results to be retrieved in one of the query types were obtained. For this, the records which should be returned by a query were previously selected, generating a baseline that was used to compare the results obtained from each approach. Once generated, the baseline was stored in a file. Next, we present the criteria used for generation of the baseline for each query type:

- **Purely spatial query:** the baseline for the feature type level was composed of all the layers whose bounding-box intersected the geographic region defined in the query. On the other hand, for the service level, the baseline was composed of all the services that offered at least one layer present in the baseline for the feature type level;

TABLE I - TABLE OF FORMALIZED NULL HYPOTHESES

Level	Canadian SDI	GEOSS SDI
Feature types	• $H_0 1 : PFT_{sd}(e) \leq PFT_c(e)$ • $H_0 2 : PFT_{sd}(t) \leq PFT_c(t)$ • $H_0 3 : PFT_{sd}(s) \leq PFT_c(s)$ • $H_0 4 : PFT_{sd}(g) \leq PFT_c(g)$ • $H_0 5 : CF_{sd}(e) \leq CF_c(e)$ • $H_0 6 : CF_{sd}(t) \leq CF_c(t)$ • $H_0 7 : CF_{sd}(s) \leq CF_c(s)$ • $H_0 8 : CF_{sd}(g) \leq CFT_c(g)$	• $H_0 17 : PFT_{sd}(e) \leq PFT_g(e)$ • $H_0 18 : PFT_{sd}(t) \leq PFT_g(t)$ • $H_0 19 : PFT_{sd}(s) \leq PFT_g(s)$ • $H_0 20 : PFT_{sd}(g) \leq PFT_g(g)$ • $H_0 21 : CF_{sd}(e) \leq CF_g(e)$ • $H_0 22 : CF_{sd}(t) \leq CF_g(t)$ • $H_0 23 : CF_{sd}(s) \leq CF_g(s)$ • $H_0 24 : CF_{sd}(g) \leq CF_g(g)$
Service	• $H_0 9 : PS_{sd}(e) \leq PS_c(e)$ • $H_0 10 : PS_{sd}(t) \leq PS_c(t)$ • $H_0 11 : PS_{sd}(s) \leq PS_c(s)$ • $H_0 12 : PS_{sd}(g) \leq PS_c(g)$ • $H_0 13 : CS_{sd}(e) \leq CS_c(e)$ • $H_0 14 : CS_{sd}(t) \leq CS_c(t)$ • $H_0 15 : CS_{sd}(s) \leq CS_c(s)$ • $H_0 16 : CS_{sd}(g) \leq CS_c(g)$	• $H_0 25 : PS_{sd}(e) \leq PS_g(e)$ • $H_0 26 : PS_{sd}(t) \leq PS_g(t)$ • $H_0 27 : PS_{sd}(s) \leq PS_g(s)$ • $H_0 28 : PS_{sd}(g) \leq PS_g(g)$ • $H_0 29 : CS_{sd}(e) \leq CS_g(e)$ • $H_0 30 : CS_{sd}(t) \leq CS_g(t)$ • $H_0 31 : CS_{sd}(s) \leq CS_g(s)$ • $H_0 32 : CS_{sd}(g) \leq CS_g(g)$

- **Purely temporal query:** for the feature type level, the baseline was composed of all the layers whose temporal extension intersected the time interval defined in the query. On the other hand, for the service level, the baseline was composed of all the services that offered at least one layer present in the baseline for the feature type level;
- **Purely semantic query:** for the feature type level, the baseline was composed of all the layers that offered data about the theme used in the request, or a theme related to it. On the other hand, for the service level, the baseline was composed of all the services that offered at least one layer present in the baseline for the feature type level;
- **Global query:** for the feature type level, we considered all the layers that met the three constraints (spatial, temporal, and semantic) defined in the request. On the other hand, for the service level, the baseline was composed of all the services that offered at least one layer present in the baseline for the feature type level.

The second step was the execution of the queries. In this stage, the query was formulated according to its criteria and performed both in SESDI and in the catalog service of the SDI.

Finally, in the last step, we compared the results obtained from the two approaches. During this stage, we generated an output file containing information about the experiment. This file stored information about the number of services and feature types in each baseline. Moreover, for each approach, the file stored the number of services retrieved, the number of feature types retrieved, the number of relevant services retrieved and the number of relevant feature types retrieved. Then, we used the information stored in that file for the computation of the precision and recall metrics.

In order to make the analysis of results easier, a tool was developed to compute the data concerning the dependent variables (precision and recall) of this research, consolidating them into a single file. This way, the consolidated data can be observed in the files related to the comparison between SESDI and the Canadian SDI [22] and between SESDI and the GEOSS SDI [23].

V. Results of the Experimental Evaluation

After collecting the data, it was necessary to employ the hypotheses tests to infer about the hypotheses cited in previous section. To accomplish this task, we used a methodology proposed by Wohlin [24]. According to this method, we firstly verified the normality, aiming to check whether the data come from a normal distribution. Next, we collected the homoscedasticity of the data in order to tell whether there is variance or not in the data. Both normality and homoscedasticity allow one to decide about the use of parametric or non-parametric tests. Since every statistical test needs a null hypothesis to infer about it, the tests used here result in a p-value (probability value) that, depending on its value, may deny or not the hypothesis of the test according to the significance level adopted.

In order to verify the normality, we used the Shapiro-Wilk, Anderson-Darling, Skewness and Kurtosis tests [24], with significance level of five percent, besides the QQ-plot graphics. Moreover, in the comparison of SESDI with the Canadian SDI and the GEOSS, the data are not from a normal distribution. Since most tests point that the data involving controlled variables with the dependent variables are not normal, a few data were pointed as normal in some tests, but the QQ-plot graphic concerning them allowed us to see their non-normality. The results of the tests and the QQ-plot graphics can be seen in [25].

Since the collected data are not normal, the Levene test [24] was used to perform the verification of the homoscedasticity aspects with a significance level of five percent. The results of the tests related to the data of the comparisons between the SESDI and the catalog services of the Canadian SDI and of the GEOSS can be seen in [25].

As we mentioned before, the data were not retrieved from a normal population. So, we used a non-parametric test to infer about the experiment hypotheses. Since the data were collected in a dependent manner, we needed to use a non-parametric test of the paired samples. Besides, the hypotheses of the experiment compared to data groups. For this reason, we chose the Mann–Whitney with significance level of five percent.

The results of the test for each hypothesis listed in Table I, related to the performance comparison between SESDI and the Canadian SDI, are presented in Table II. Such results are coded in two colors: green, when the null hypothesis was not refuted (p-value is above the significance level adopted - 5%) and red, when the null hypothesis was refuted (p-value is smaller than the significance level adopted - 5%). Table II shows that most of the hypotheses related to comparison between SESDI and the Canadian SDI were refuted. These results lead us to conclude that SESDI had better performance for most cases, since the null hypotheses presented assume the inferiority or equality of the response variables for all query types. The table shows that the hypotheses *H0-5* and *H0-13* were not refuted, which means that we cannot state that SESDI improves the recall for spatial queries at service and feature type levels.

Table II - Results of the Mann-Whitney Tests Applied to the Comparison of SESDI with the Canadian SDI

RESPONSE VARIABLE	HYPOTHESIS	RESULT
Precision at feature types level	H0-1	p-value = 6.77e-05
	H0-2	p-value = 3.585e-05
	H0-3	p-value = 0.0001568
	H0-4	p-value = 0.0004186
Coverage at feature types level	H0-5	p-value = 1
	H0-6	p-value = 0.0003208
	H0-7	p-value = 2.645e-05
	H0-8	p-value = 0.001244
Precision at service level	H0-9	p-value = 0.0002134
	H0-10	p-value = 3.585e-05
	H0-11	p-value = 0.009584
	H0-12	p-value = 0.0005406
Coverage at service level	H0-13	p-value = 1
	H0-14	p-value = 2.645e-05
	H0-15	p-value = 9.947e-05
	H0-16	p-value = 0.0005439

The results of the test for each hypothesis listed in Table I, related to the performance comparison between the SESDI and the GEOSS SDI, are presented in Table III. In that table, it is possible to notice also that most of the hypotheses related to the performance comparison between the SESDI and the GEOSS SDI were refuted, which allows us to conclude that the performance of the SESDI was superior to that of the GEOSS SDI for most cases. The only hypothesis that was not refuted (*H0-22*) shows that there is no statistical significance to state that SESDI improved recall of temporal queries at feature type level.

To better illustrate the results presented above, the returns a semantic query used in the experiment are shown below. The consultation aimed to find maps for "boundaries" and obtained a precision level of feature type 100% in the SESDI, i.e., all relevant services were recovered, while the IDE obtained at a precision level of 11% of feature type.

Table III - Results of the Mann-Whitney Tests Applied to the Comparison between SESDI and GEOSS

RESPONSE VARIABLE	HYPOTHESIS	RESULT
Precision at feature types level	H0-17	p-value = 2.725e-05
	H0-18	p-value = 1.576e-05
	H0-19	p-value = 0.0001048
	H0-20	p-value = 9.23e-05
Coverage at feature types level	H0-21	p-value = 9.666e-06
	H0-22	p-value = 0.1072
	H0-23	p-value = 1.576e-05
	H0-24	p-value = 0.001668
Precision at service level	H0-25	p-value = 2.543e-05
	H0-26	p-value = 1.265e-05
	H0-27	p-value = 0.0001438
	H0-28	p-value = 0.0001146
Coverage at service level	H0-29	p-value = 3.018e-05
	H0-30	p-value = 1.576e-05
	H0-31	p-value = 0.001187
	H0-32	p-value = 0.0004807

VI. Conclusion and Future Works

This research was intended to perform a deeper evaluation of the performance of the SESDI framework during the retrieval of geographic data offered by SDIs. In order to perform this evaluation, the performance of SESDI was compared with the performance of two catalog services presently offered by two infrastructures: The Canadian SDI and the GEOSS SDI. Faced with the results presented above, one can conclude, with a significance level of five percent, that:

1) For all types of queries, SESDI had better precision than the Canadian SDI for most of the queries executed at the feature type level;
2) For the global, purely temporal and purely semantic query types, SESDI had a better recall at feature type level, compared to the Canadian SDI. As for the purely spatial queries, there was not statistical significance to state which approach had better performance;
3) For all types of query, SESDI had better precision in the service level than the Canadian SDI;
4) For the global, purely temporal and purely semantic query types, SESDI had a better recall at service level than the Canadian SDI. As for the purely spatial queries, there was not statistical significance to assert which approach had better performance;
5) For all types of query, SESDI had better precision at the feature type level than the GEOSS SDI;
6) For the purely spatial, purely semantic and global query types, SESDI had a better recall at the feature type level than the GEOSS SDI. As for the purely spatial queries, there was not statistical significance to state which approach had better performance;
7) For all types of query, SESDI had better precision at service level than the GEOSS SDI;
8) For all types of query, SESDI had better recall at service level than the GEOSS SDI.

Based on the results, we could conclude that the solution used to implement SESDI is viable, since it improved recall and precision for most queries used during experimental evaluation. As a suggestion for future work, SESDI could be compared with other catalog services. The results achieved with such experiments could add still more value to the object of study of this research.

Another suggestion for future works would be the implementation of the universalization of the language used by SESDI. With this, it would be possible to compare SDIs with languages other than English.

References

[1] Lupp, M. (2008). Open geospatial consortium. In Shekhar, S. and Xiong, H., editors, Encyclopedia of GIS, page 815. Springer.

[2] Open Geospatial Consortium, “Ogc web map service interface,” OGC 03-109r1, Version 1.3.0, January 2004.

[3] ——, “Web feature service implementation specification,” OGC 04-094, Version 1.1.0, May 2005.

[4] F. G. de Andrade, C. de S. Baptista, and C. A. Davis Jr, “Improving geographic information retrieval in spatial data infrastructures.” GeoInformatica, Jan. 2014, doi: 10.1007/s10707-014-0202-x.

[5] http:// http://www.fgdc.gov/nsdi/nsdi.html [accessed: 2014-01-20]

[6] http://geoconnections.nrcan.gc.ca/[accessed: 2014-01-20]

[7] http://www.earthobservations.org/geoss.shtml[accessed: 2014-01-20]

[8] K. M. Stock et al., “A semantic registry using a feature type catalogue instead of ontologies to support spatial data infrastructures.” International Journal of Geographical Information Science, vol. 24, no. 2, Feb. 2010, pp. 231–252.

[9] M. Lutz, J. Sprado, E. Klien, C. Schubert, and I. Christ, “Overcomingsemantic heterogeneity in spatial data infrastructures.” Computers Geosciences, vol. 35, no. 4, Apr. 2009, pp. 739–752.

[10] M. Lutz, “Ontology-based descriptions for semantic discovery and composition of geoprocessing services.” GeoInformatica, vol. 11, no. 1, Mar. 2007, pp. 1–36.

[11] R. Lemmens, R. A. de By, M. Gould, A. Wytzisk, C. Granell, and P. van Oosterom, “Enhancing geo-service chaining through deep service descriptions,” T. GIS, vol. 11, no. 6, Dec. 2007, pp. 849–871.

[12] M. Molina and S. Bayarri, “A multinational sdi-based system to facilitate disaster risk management in the andean community.” Computers Geosciences, vol. 37, no. 9, Sep. 2011, pp. 1501–1510.

[13] N. Wiegand and C. García, “A task-based ontology approach to automate geospatial data retrieval,” T. GIS, vol. 11, no. 3, Jun. 2007, pp. 355–376.

[14] W. Li, et al., “Semantic-based web service discovery and chaining for building an arctic spatial data infrastructure.” Computers Geosciences, vol. 37, no. 11, Nov. 2011, pp. 1752–1762.

[15] K. Janowicz, M. Wilkes, and M. Lutz, “Similarity-based information retrieval and its role within spatial data infrastructures.” in GIScience, ser. Lecture Notes in Computer Science, T. J. Cova, H. J. Miller, K. Beard, A. U. Frank, and M. F. Goodchild, Eds., vol. 5266. Springer, Sep. 2008, pp. 151–167.

[16] C. Zhang, T. Zhao, W. Li, and J. P. Osleeb, “Towards logic-based geospatial feature discovery and integration using web feature service and geospatial semantic web.” International Journal of Geographical Information Science, vol. 24, no. 6, Jun. 2010, pp. 903–923.

[17] Z. Li, C. P. Yang, H. Wu, W. Li, and L. Miao, “An optimized framework for seamlessly integrating ogc web services to support geospatial sciences.” International Journal of Geographical Information Science, vol. 25, no. 4, Apr. 2011, pp. 595–613.

[18] P. C. Smits and A. Friis-Christensen, “Resource discovery in a european spatial data infrastructure.” IEEE Trans. Knowl. Data Eng., vol. 19, no. 1, Jan. 2007, pp. 85–95.

[19] N. Athanasis, K. Kalabokidis, M. Vaitis, and N. Soulakellis, “Towards a semantics-based approach in the development of geographic portals,” Comp. Geosci, vol. 35, Feb. 2009, pp. 301–308.

[20] N. Chen, Z. Chen, C. Hu, and L. Di, “A capability matching and ontology reasoning method for high precision ogc web service discovery.” Int. J. Digital Earth, vol. 4, no. 6, Jun. 2011, pp. 449–470.

[21] http://zip.net/bcmbc2[accessed: 2014-01-20]

[22] http://zip.net/bkmbzd[accessed: 2014-01-20]

[23] http://zip.net/bkmbzc[accessed: 2014-01-20]

[24] C. Wohlin et al., Experimentation in Software Engineering. Heidelberg, Berlin: Springer-Verlag, 2012.

[25] http://zip.net/bll9Wz[accessed: 2014-01-20]