INDE METADATA CONFORMITY INDICATOR

Taís Virgínia Gottardo[1] – ORCID: 0000-0003-4234-8569

Ivanildo Barbosa[2] – ORCID: 0000-0003-4164-0875

[1] Instituto Militar de Engenharia, Programa de Pós-Graduação em Engenharia de Defesa,
Rio de Janeiro, Rio de Janeiro, Brasil.
E-mail: tvgottardo@gmail.com

[2] Instituto Militar de Engenharia, Departamento de Engenharia Cartográfica,
Rio de Janeiro, Rio de Janeiro, Brasil.
E-mail: ivanildo@ime.eb.br

Abstract:

Metadata represents a set of descriptive information about the data, which aims at facilitating the search, access and use of data. The metadata standards specify the minimum set of elements to be informed and the file structure to ensure interoperability among catalogs. The present work aims at analyzing the adherence of the metadata published in the Brazilian Spatial Data Infrastructure (SDI) repository to the MGB Profile (in Portuguese, Perfil de Metadados Geoespaciais do Brasil) through a conformity indicator. Conformity was evaluated for over 30,000 metadata inserted in the INDE catalog (in Portuguese, Infraestrutura Nacional de Dados Espaciais), in June 2018, considering the set of mandatory and conditional elements of the summarized MGB Profile. These elements were collected from a harvester elaborated by the authors capable of scanning all the metadata of the repository and returning the fields organized into a CSV file. After analysis, it was found that only 28% of metadata conforms to the summarized MGB Profile. The low conformity rate suggests the limited understanding of the standards recommended, regarding both the minimum elements to be informed about each product and the structuration of the information in XML.

Keywords: spatial metadata; SDI; ISO 19115; XML; metadata quality; MGB Profile

1. Introduction

Spatial Data Infrastructures (SDIs) are mechanisms of standardization, governance, dissemination and access to geospatial data. To fulfill this role, SDIs adopt norms and standards that enable interoperability of data and metadata, which are of crucial importance to make applications of interest to government and society (CONCAR, 2010).

> Catalog services and registries, with their underlying catalogs or metadata repositories, are a fundamental part of any SDI. The catalogs allow clients and services to discover resources and to evaluate if they are fit for purpose. (Smits and Friis-Christensen, 2007, p. 85)

Metadata standards are a set of elements necessary to characterize the data produced by an institution. ISO 19115-1 is the main geospatial metadata standard, determining an extensive list of metadata elements, defining a minimum set of metadata that should be common to any profile based on this standard, and defining the rules for profile creation (ISO, 2014). ISO 19139 confers the metadata implementation rules in XML format, structuring the data according to appropriate vocabulary (ISO, 2007). In Brazil, the MGB Profile (in Portuguese, Perfil de Metadados Geoespaciais do Brasil) specifies the metadata standard developed by the CONCAR (in Portuguese, Comissão Nacional de Cartografia) and adopted by the INDE (in Portuguese, Infraestrutura Nacional de Dados Espaciais).

Geospatial metadata documents information about geospatial data such as identification, reference system, content information, quality, and distribution. Despite the importance given to the production of metadata, there is still significant resistance to its documentation, which is commonly considered a monotonous and time-consuming task (Olfat et al., 2010), as well as problems with standards understanding that are in general extensive and complex (Manso-Callejo et al., 2010), factors that contribute to the occurrence of failures and omissions in completing the information.

A questionnaire was applied to cataloging, creation and maintenance of metadata professionals, indicating that approximately 83% and 68% of the respondents emphasized, respectively, the importance of the content standards and the semantics of the metadata, as critical factors of metadata quality (Park and Tosaka, 2010).

Nonconformity may imply inefficient use of the resources for both search and retrieval of geospatial products through the INDE metadata catalog (INDE, 2018a), as well as limiting the reuse of the metadata produced. On the other hand, the conformity to the metadata standard promotes interoperability between data repositories, making possible the reuse of a metadata in different repositories or the catalogs unification. Standard conformance also facilitates metadata upgrade when profiles are reviewed. The properly documented metadata maintenance can be automated or semi-automated.

The objective of the present work is analyzing the adherence of the metadata published in the INDE repository to the Brazilian Geospatial Metadata Profile - MGB Profile (CONCAR, 2009) through a conformity indicator.

This paper is organized as follows: Section 2 presents a literature review of geospatial metadata standards and previous work about metadata quality to be considered for the current analysis. Section 3 describes the adopted criteria to evaluate the metadata conformity and the methodology elaborated for such analysis. Section 4 presents the results, the 'general indicator' and the analysis for metadata element to guide producers to the most critical issues. Finally, conclusions, final considerations about the elaborated methodology, and suggestions are related in section 5.

2. Related works on metadata quality and metadata standards

2.1 Metadata Quality

Metadata is commonly defined as data about data. Nowadays, the availability of online data constantly increases, and metadata are stored and distributed by repositories or catalogs. These digital collections have functions like promoting discovery, identification, selection, and use of digital resources for the interested community and the metadata quality is crucial for building good digital collections (Park and Tosaka, 2010). Tani et al. (2013, p. 1194) affirm that metadata quality "deeply influences the overall quality of the services offered by relying on the data they characterize". Park and Tosaka (2010) emphasize that lack of quality in metadata records depreciates the value of digital collections for potential users.

Approaches to metadata quality are diverse. Bruce and Hillmann (2004) relate seven general and domain-independent characteristics of metadata quality: completeness, accuracy, provenance, conformance to expectations, logical consistency and coherence, timeliness, and accessibility. Ochoa and Duval (2009, apud Tani et al., 2013) complement Bruce and Hillmann framework and propose 13 quality metrics to evaluate the metadata quality of a collection. Stvilia et al. (2004) analyze information quality dimensions and determine more than 20 metadata quality parameters, classified as intrinsic, relational and reputational. NISO (2007) lists six principles for the creation of good metadata, they are: conformity to community standards, interoperability, authority control and use of content standards to describe objects and collocate related objects, clear statement of the conditions and terms of use, cover the long-term curation and preservation of objects in collections, existence of the qualities of good objects (authority, authenticity, archivability, persistence, and unique identification). Park (2009) presents a state of art in metadata quality and point out that accuracy, completeness, and consistency are the most commonly used criteria in assessment metadata quality.

The most relevant studies in metadata quality are based on library and information science knowledge, but they generally are generics enough to be adapted to a variety of applications. The tendency of different expert communities to resist generic solutions and seek a unique solution to general problems can create barriers to evolution in metadata quality science (Bruce and Hillmann, 2004).

In the present paper, metadata quality is assessed in the Spatial Data Infrastructures (SDIs) context. The quality parameter considered was conformity to the metadata standard adopted, i.e. conformance to the community expectations.

Many researches explore and analyze the quality of metadata stored in repositories under different perspectives (Bui and Park, 2006; Díaz et al., 2012; Rousidis et al., 2014; Balatsoukas et al., 2018). In some of them (Bui and Park, 2006; Rousidis et al., 2014; Balatsoukas et al., 2018), the harvesting of the repository metadata is done using the OAI-PMH protocol (Open Archives Initiative Protocol for Metadata Harvesting) for the following analysis. The OAI-PMH Validator & Data extractor is able to download all records from digital libraries, parallelly or individually, and analyze compliance of metadata with OAI-PMH, Dublin Core (DC) and other standards (Banos, 2011). The data harvesting methodologies used by Bui and Park (2006), Rousidis et al. (2014) and Balatsoukas et al. (2018) are similar and follow the steps: download of all metadata from catalog, format transformation (involving XML merge) and data analysis using popular tools like Excel® or Access®.

Despite advantages of this tool, the implementations of OIA-PMH are usually limited to Dublin Core metadata that are simple metadata content standard, opposite to complex metadata standard as ISO 19115 (Schindler and Diepenbroek, 2008). Like Schindler and Diepenbroek (2008) this work also present an own methodology for harvesting documents from CSW (Catalog Service for the Web).

The methodology is detailed in section 3. The following topics present useful references for this work, a literature review of spatial metadata standards which are relevant to Brazilian geospatial information producers, as well as similar studies on metadata conformity. It should be noted that academic works regarding geospatial metadata quality is scarce in opposition to generic metadata research.

2.2 Metadata ISO standards and MGB Profile

ISO 19115 is a standard developed by the ISO/TC 211 - Geographic Information/Geomatics Technical Committee that establishes a standard for the generation and organization of geospatial metadata. This committee deals with standardization in the field of digital geographic information and is currently constituted by 37 participating countries and 27 observer countries.

In the elaboration of this standard, standards already defined by other organizations such as the FGDC (Federal Geographic Data Committee), ANZLIC (Australia and New Zealand Land Information Council) and CEN (European Committee for Standardization) were analyzed (Freitas, 2005).

With information of title, reference date, language, abstract, distribution information, among others, the metadata constructed based on this standard have the purpose of facilitating the management and organization of geospatial data, making the use of the data more efficient, facilitating the location, access, evaluation and use of the data, allowing users to verify which data is best suited to their applications. Based on ISO 19115, producers describe their data with appropriate information.

ISO 19115-1: 2014 presents a framework for describing digital geospatial data through metadata (ISO, 2014), defining:

- Mandatory and conditional metadata sections, entities, and metadata elements;

- Minimum set of metadata needed to meet a range of applications - called core;

- Normative instruction for elaboration of extensions and profiles.

ISO 19139: 2007 defines an implementation scheme in XML for ISO 19115, that is, it standardizes the grammar of the digital files of geospatial metadata. The XML (eXtensible Markup Language) language allows the configuration of elements in tags, which can be structured in a hierarchical way, emphasizing the great interoperability.

In Brazil, the adaptation of ISO 19115:2003 (ISO, 2003) to its cartographic needs resulted in the publication of the MGB Profile. This document was approved by the CONCAR in November 2009 and should be adopted by the INDE actors.

The current version of the MGB Profile was prepared based on ISO 19115 of 2003 and has not yet been revised considering the updates of ISO 19115-1: 2014. It consists of a selection of elements presented in the ISO that form a set which is capable of describing the data produced by the main agents of the Brazilian cartography.

In the standard proposed for Brazil there is no inclusion of new packages, classes or elements, only a subtle reorganization of sections, classes and elements.

The information packages covered by the MGB Profile are: identification, data identification, constraint information, data quality, maintenance information, spatial representation information, reference system, content information, distribution and metadata on metadata (CONCAR, 2009).

A subset of MGB Profile elements, containing the minimal components required to describe geospatial data is named summarized MGB Profile. The summarized version of the profile includes: mandatory elements, which must be present in all metadata produced based on the MGB Profile; conditional elements, which must be present in all metadata that meet a given condition; and optional elements that may or may not be considered in a metadata produced based on this standard, but which, being contained in the core, are strongly recommended elements (CONCAR, 2009).

The mandatory metadata elements of the summarized MGB Profile are: title, date, responsible party, language, topic category, abstract, distribution format, reference system, metadata responsible party, metadata date and status. And the conditional metadata elements of the summarized MGB Profile are: geographic extension, geospatial data set character code, metadata language, and metadata character code (CONCAR, 2009).

According to the CONCAR (2009), the summarized version of the profile should be adopted in case the organization does not have sufficient elements to complete the full version of this profile.

2.3 Metadata conformity of other repositories

Bui and Park (2006) analyzed the filling of 15 Dublin Core elements for metadata records from the NSDL repository. They found that the six elements indicated as more critical for searching and retrieval purposes are commonly well populated, but there are neglected elements. Examples

of percentage of non-empty records per element: Title (~100%), Identifier (~99%), Creator (~83%), Descriptor (~83%), Subject (~77%), Type (~75%), Contributor (~9%), Relation (~7%) and Coverage (~2%).

INSPIRE (Infrastructure for Spatial Information in Europe), the spatial data infrastructure of European Union members, adopts an indicator that quantifies the metadata conformity with the metadata implementation rules of the INSPIRE directive for data sets and geographic data services (MDi2) (European Commission, 2009). The INSPIRE directive also requires member states to annually report on the compliance level of their metadata.

Figure 1 represents the metadata conformity results obtained for INSPIRE member states in 2017, except for French Guiana that does not appear in this figure. The value 1 represents 100% of metadata conformity. In this map it is possible to observe that there are not values below 0,34 (that is, 34%) and most of the member states present conformity close to 100%.
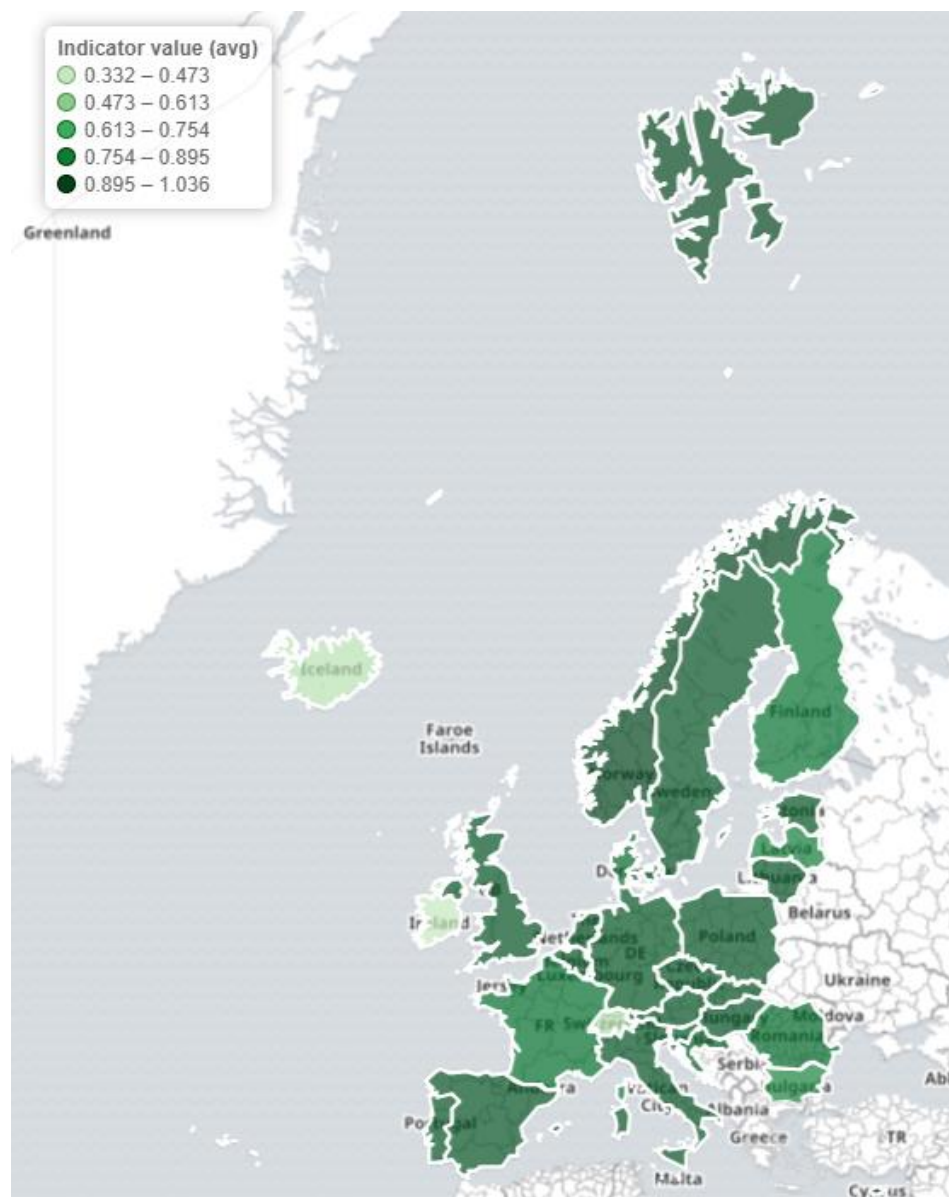

Figure 1: Metadata conformity results for INSPIRE member states – 2017 (INSPIRE, 2018).

Source: https://inspire-dashboard.eea.europa.eu

2.4 Geospatial metadata conformity of Brazilian producers

Pascoal et al. (2013) compared metadata files with XML schemas for the complete MGB Profile and summarized MGB Profile, performing the validation of 50 XML metadata files obtained in different catalogs available online.

As a result, they found that none of the 50 files analyzed were fully conformed with either of the two verification schemes considered. The main errors were: element not allowed, when, for example, the model schema predicts that an element is a real type, but in the XML file it is implemented as a field of character string type; elements not defined, when elements (tags) not found in the template schema are found in the XML file; and non-defined attributes, when in the XML file attributes not found in the template schema are found.

The results showed the need to adapt the metadata available in XML structure to the models adopted by the CONCAR.

As a solution, Pascoal et al. (2013) proposed materialization and availability of complete MGB Profile schemas and summarized MGB Profile in derived XML schemas from ISO 19139: 2007 to the producer institutions.

Since 2014, templates based on the MGB Profile have been available for download in the 'Help' tab of the INDE metadata catalog (INDEb, 2018) in order to help geospatial information producers. So they would not need to carry out the implementation of the MGB Profile in XML to insert their metadata into the metadata catalog.

Three different templates are available. The first one corresponds to the implementation of the complete MGB Profile for data in vector format; the second corresponds to the complete MGB Profile for data in raster format, and the third corresponds to the summarized MGB Profile.

Loti et al. (2017) analyzed the conformity of these three XML metadata models available in the INDE geoportal, comparing them with the specifications defined in the MGB Profile. They observed inconsistencies in the three available templates comparing them to the homologated profile. For the template of summarized MGB Profile excesses of elements and divergences from the specification were observed. In this case, it was verified that 13 elements of the template were not present in the specification (i. e. there were more elements) and five elements were in a format which differs from that predicted in profile. Therefore, it is possible to infer that the templates provided in the INDE geoportal do not fully comply with the CONCAR standards.

3. Metadata conformity indicator

3.1 Metadata conformity

In this research, metadata inserted in the INDE metadata catalog are considered conform if they:

(a) include the mandatory elements of the summarized MGB Profile;

(b) include the conditional elements of the summarized MGB Profile, if the corresponding products meet the proposed condition;

(c) present the elements listed in items (a) and (b) structured in XML following the structuring provided by ISO 19139: 2007.

The summarized MGB Profile consists of 23 metadata elements, 11 of which are mandatory, i.e. meet the premise (a) (Table 1); 4 are conditional, i.e. meet the premise (b) (Table 2); and 8 are optional. Thus, a set of 15 metadata elements was investigated.

The MGB Profile, as well as ISO 19139: 2007, specifies the tags where the elements must be inserted and the element data type, be it string, date format, codelist or others.

For the 15 elements to be analyzed, the tags to be consulted, as well as their hierarchy, were identified. Therefore, tags structured in disagreement with the planned XML structuring were not considered valid and were also counted as cases of non-conformity. Table 1 and Table 2 show elements, tags and data type for mandatory and conditional elements, respectively.

Table 1: Mandatory elements, tags and data type

Element	ISO 19115:2003 Context	Data Type
Title	MD_Metadata.identificationInfo>MD_Identification.citation>CI_Citation.title	Character string
Date	MD_Metadata.identificationInfo>MD_DataIdentification.citation>CI_Citation.date	Date
Responsible party (individual name or organisation name)	MD_Metadata.identificationInfo>MD_Identification.pointOfContact>CI_ResponsibleParty.individualName or MD_Metadata.identificationInfo>MD_Identification.pointOfContact>CI_ResponsibleParty.organisationName	Character string
Language	MD_Metadata.identificationInfo>MD_Identification.language	codelist
Topic category	MD_Metadata.identificationInfo>MD_DataIdentification.topicCategory	codelist
Abstract	MD_Metadata.identificationInfo>MD_Identification.abstract	Character string

Table 1: Cont.

Distribution format (searched by format name)	MD_Metadata.distributionInfo>MD_Distribution.distrib utionFormat>MD_Format .name	Character string
Reference system	MD_Metadata.referenceSystemInfo>MD_ReferenceSyst em.referenceSystemIdentifier>RS_Identifier	Character string
Metadata responsible party (individual name or organisation name)	MD_Metadata.contact>CI_ResponsibleParty.individualN ame or MD_Metadata.contact>CI_ResponsibleParty.organisatio nName	Character string
Metadata date	MD_Metadata.dateStamp	date
Status	MD_Metadata.identificationInfo>MD_Identification.stat us>MD_ProgressCode	codelist

Table 2: Conditional elements, tags and data type

Element	ISO 19115:2003 Context	Data Type
Geographic extension	MD_Metadata.identificationInfo>MD_DataIdentification. extent>EX_Extent.geographicExtent>EX_BoundingBox.we stBoundLongitude and MD_Metadata.identificationInfo>MD_DataIdentification. extent>EX_Extent.geographicExtent>EX_BoundingBox.eas tBoundLongitude and MD_Metadata.identificationInfo>MD_DataIdentification. extent>EX_Extent.geographicExtent>EX_BoundingBox.so uthBoundLatitude and MD_Metadata.identificationInfo>MD_DataIdentification. extent>EX_Extent.geographicExtent>EX_BoundingBox.no rthBoundLatitude	angle
Geospatial data set character code	MD_Metadata.identificationInfo>MD_DataIdentification. characterSet	codelist
Metadata language	MD_Metadata.language	codelist
Metadata character code	MD_Metadata.characterSet>MD_characterSetCode	codelist

3.2 Harvesting and analyzing of metadata records

In order to analyze the 35,260 XML metadata files available in the INDE metadata catalog on June 2, 2018, a Node.js script was elaborated to extract the necessary information and compose a database in Comma Separated Values (CSV) format, allowing analysis in spreadsheets. Node.js is a JavaScript runtime built on Chrome's V8 JavaScript engine (Node.js Foundation, 2018). Node.js is a collaborative and open source project.

The harvesting script elaborated for this work searches the tags of interest in the repository and therefore has advantages over other harvesters. Unlike the OIA-PMH protocol, the elaborated harvester does not download the complete metadata file from the repository; instead, it harvests only the 23 tags of interest of this work and returns a CSV file as output, also eliminating the need for format transformations. For comparison, in this work were collected elements of 35,260 XML files from the repository generating an output 3.4MB file; Balatsoukas et al. (2018) downloaded 135MB from a total of 516 xml files.

The harvester operates in two steps. First, the file identifiers of metadata records are harvested by recursive request and an ID list is composed. Second, each metadata record is requested by its identifier and the tags are collected.

For elements with data type character string, the only analysis carried out was if the field was filled out (true/false test, being 'true' to filled, 'false' to not filled). For elements with data type codelist, such as the status element, it was evaluated whether the field was filled with some value of the list or not (true/false test, being 'true' for a codelist value, 'false' for non-codelist value).

Metadata conformity could be analyzed from spreadsheets in which the columns indicated the metadata elements investigated and the rows indicated the results of the true/false test.

In the third step, only the mandatory elements conformity were analyzed. In the fourth step, the conditional elements of the set of metadata considered conform in the third step were analyzed. Figure 2 shows a framework of metadata conformity methodology for the INDE metadata catalog, the harvesting process (step 1 and step 2) and the data analysis (step 3 and step 4).
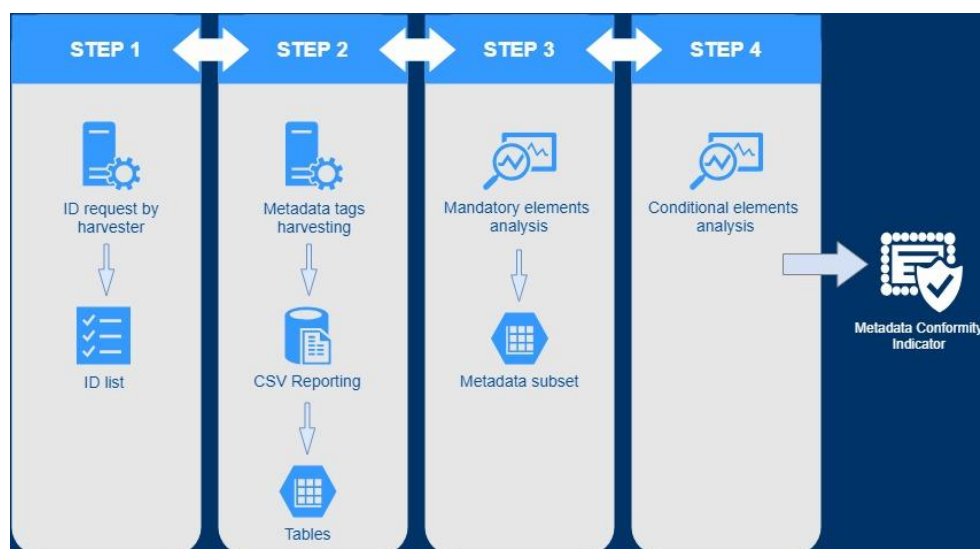

Figure 2: Harvesting and analysis steps.

4. Results

4.1 INDE Metadata Conformity Indicator

The first two steps were performed flawlessly. The conformity analysis of mandatory elements resulted that about 10,000 cataloged metadata the met premise (a) (step 3 in Figure 2).

For the analysis of conditional elements - character code, geographical extension, metadata language and data character code, the metadata character code element was conformed for all metadata. About the other three elements, only 90 products presented some kind of issue in the filling of theses conditional elements.

These 90 cases were accessed one by one to verify whether the set of data described by the metadata met the necessary conditions for these elements. In case to meet the conditions, the empty element was considered non-conform. This analysis resulted that these 90 cases were non-conform.

In another situation, a much larger amount of metadata with conditional elements without filling is possible. To facilitate the investigation of these elements it is suggested to test additional elements. In the case of geographic extension, for example, filling is mandatory if there is no documentation of altimetric-bathymetric extension or temporal extension. Thus, for automated analysis of this element, it would be possible to test the three fields (geographic extent, altimetric-bathymetric extension and temporal extension), and for conformity, one of the three fields should be filled.

The conformity index found for the totality of the metadata available to users through the INDE metadata catalog, on June 2, 2018, considering the presented methodology, was 28%. This means that only 28% of the available metadata met premises (a), (b) and (c).

4.2 Conformity for each metadata element

Table 3 presents the calculated conformity indicator for each metadata element. The best results were found for the elements: title, geographical extension, metadata responsible party and metadata date. These elements showed conformity in almost 100% of the analyzed XML files. The worst results were given for the elements: distribution format (66% of non-conformity cases), data character code, status and topic category.

In cases of non-conformity of the element of distribution format, it is understood that there are non-distributed data (restricted cases, for example) or data distributed only in physical format (for example printed maps). However, if it is a mandatory element, the field should be filled with some value. One possible solution would be to fill in the information 'not applicable'.

Table 3: Conformity indicator for each metadata element

Element	Conform	Non-conform
Title	99,8%	0%
Geographic extension	99,7%	0%
Metadata responsible party	99%	1%
Metadata date	99%	1%
Metadata language	87%	13%
Abstract	45%	55%
Responsible party	45%	55%
Date	44%	56%
Metadata character code	44%	56%
Language	43%	57%
Reference system	38%	62%
Topic category	37%	63%
Status	37%	63%
Geospatial data set character code	37%	63%
Distribution format	34%	66%

5. Conclusions

The aim of the present work was analyzing the adherence of the metadata published in the INDE repositories to the MGB Profile through a conformity indicator. To achieve the proposed results, a method was implemented to harvest the metadata elements from the INDE repository and to analyze the conformity. Among other implications, non-conformity compromises catalog interoperability and makes multi-value searches impossible.

The methodology elaborated for the harvesting of the metadata elements of interest presented advantages over other harvesters, as discussed in section 3.2. With few changes, the elaborated harvester can be adapted to provide the compliance indicator semi-automatically. According to the definition of conformity adopted in this work, it is possible to automatically analyze the conformity of mandatory elements. For conditional elements, semi-automatic solutions can be adopted (see an example of solution in section 4.1), eliminating considerably a technician work.

In addition, the elaborated harvester can also be reconfigured to collect tags other than the 15 collected. Therefore, a multitude of analysis of the repository metadata set could be made.

Regarding the value obtained for the conformity of the current metadata of the catalog, the low conformity rate found, 28%, indicates that the geospatial information producers responsible for the inclusion of the metadata in the catalog have a limited understanding of the standards recommended by the CONCAR, regarding both the minimum elements to be informed for the product and the structuring of information in XML, although the MGB Profile was published more than 10 years ago and has been available online ever since.

By analyzing the individual elements, conformity exceeded 99% for Title, Geographic extension, Metadata responsible party, and Metadata date. On the opposite side, less than 40% of values for Reference system, Topic category, Status, Geospatial data set character code and Distribution format were found conform.

The result of the indicator assists in identifying and scaling the problem. Defining the conformity for elements helps in the direction of the solutions, since it makes possible to detect which information demands more efforts for adequacy. The results of this research can, for example, assist in the automatic correction of elements. The metadata language element is an example that can be automatically populated with the use of tools that detect languages of the already filled fields.

Other analysis, such as the level of compliance of metadata per institution, could also be useful in this diagnosis.

The existence of mechanisms to both control and fix nonconformities is essential at the time of registration of the metadata or in the case of standards update. In the literature, it is possible to find many recommendations to improve the quality of metadata since its creation. Park and Tosaka (2010) lists some measures: use of mechanisms like as drop-down menus and pop-up windows with textual guidance based on metadata scheme, use of semi-automatic metadata generation tools, staff training, and periodic sampling of metadata records for quality review.

In future works, other evaluations can be done automatically or semi-automatically, such as spelling errors detection and values' format consistency. Other analysis can be done as verification of the relevance of the content or veracity of the provided information in the metadata in comparison to the data.

Besides these analysis, future works can propose indicators that evaluate the real value of the repository as a service and not only of the objects contained in it, for example, usability searches. Other quality approaches can also be evaluated as relevance of information, interoperability, timeliness, etc.

It is expected that the analyzes and discussions presented in this paper reinforce the need not only for the adoption of standards, but also for the rigorous application and maintenance, contributing to cultural change of the geographic data documentation. It is also expected that this work will contribute to the improvement of the INDE metadata catalog, which will bring benefits to the entire community interested in geospatial information.

Author´s contribution

Acknowledgement

REFERENCES

Balatsoukas, P., Rousidis, D. and Garoufallou, E. 2018. A method for examining metadata quality in open research datasets using the OAI-PMH and SQL queries: the case of the Dublin Core 'Subject' element and suggestions for user-centred metadata annotation design. *International Journal of Metadata, Semantics and Ontologies*, 13(1), pp. 1-8.

Banos, V., 2011. *OIA PMH Validator*. [online] Available at: <http://validator.oaipmh.com/>. [Accessed 30 January 2019].

Bruce, T. R. and Hillmann, D. I. 2004. *The continuum of metadata quality: defining, expressing, exploiting*. ALA editions.

Bui, Y. and Park, J.R. 2006. An assessment of metadata quality: A case study of the national science digital library metadata repository. In: *Proceedings of the Annual Conference of CAIS*. Toronto, Ontario, 1–3 June 2006.

CONCAR – Comissão Nacional de Cartografia. 2009. *Perfil de Metadados Geoespaciais do Brasil – Perfil MGB*. Brasília: Ministério do Planejamento.

CONCAR – Comissão Nacional de Cartografia. 2010. *Plano de Ação Para Implantação da Infraestrutura Nacional de Dados Espaciais*. Rio de Janeiro.

Díaz, P. et al. 2012. Analysis of quality metadata in the GEOSS Clearinghouse. *International Journal of Spatial Data Infrastructures Research*, 7, pp. 352-377.

European Commission. 2009. Commission Decision regarding INSPIRE monitoring and reporting. *Official Journal of the European Union*. L 148. pp. 18-26.

Freitas, A. L. B. 2005. *Catálogo de Metadados de Dados Cartográficos como suporte a implementação de Clearinghouse Nacional*. PhD thesis, Instituto Militar de Engenharia, Rio de Janeiro.

INDE – Infraestrutura Nacional de Dados Espaciais, 2018a. *Catálogo de Metadados*. Available at: <http://www.metadados.inde.gov.br/>. [Accessed 6 November 2018].

INDE – Infraestrutura Nacional de Dados Espaciais., 2018b. *Template Perfil MGB Sumarizado.* Available   at: <http://www.metadados.inde.gov.br/geonetwork/Download/iso19139.mgbsumarizado.zip>. [Accessed 16 October 2018].

INSPIRE, 2018. *Indicator thematic map - Ref. year 2017 - Metadata availability and conformity*. Available at: < https://inspire-dashboard.eea.europa.eu>. [Accessed 15 October 2018].

ISO – International Organization for Standardization. 2003. *Geographic Information – Metadata. ISO 19115:2003*. 1st ed. London, England.

ISO – International Organization for Standardization. 2007. *ISO 19139:2007*: *Geographic Information – Metadata – XML schema implementation.*

ISO – International Organization for Standardization, 2014. *ISO 19115-1:2014: Geographic information – Metadata – Part 1: Fundamentals*. Available at: <https://www.iso.org/standard/53798.html>. [Accessed 6 November 2018].

Loti, L. B. S. et al. 2017. Análise da Conformidade dos Templates Disponíveis na INDE com o Perfil de Metadados Geoespaciais do Brasil. In: *XXVII Congresso Brasileiro de Cartografia*. Rio de Janeiro, Brazil, 6-9 November 2017, pp. 1176-1180.

Manso-Callejo, M., Wachowicz, M. and Bernabé-Poveda, M. 2010. The design of an automated workflow for metadata generation. In: *4th international conference, MTSR 2010*. Alcalá de Henares, Spain, 20–22 October 2010, pp. 275-287.

NISO - National Information Standards Organization. 2007. *A Framework of Guidance for Building Good Digital Collections*, Baltimore, MD, 61–2.

Node.js Foundation, 2018. About Node.js®. Available at: <https://nodejs.org/en/about>. [Accessed 17 October 2018].

Olfat, H., Rajabifard, A. and Kalantari, M. 2010. A synchronisation approach to automate spatial metadata updating process. *Coordinates Magazine*, VI (3), pp. 27-32.

Park, J. R. 2009. Metadata quality in digital repositories: A survey of the current state of the art. *Cataloging & classification quarterly*, 47(3-4), pp. 213-228.

Park, J. R. and Tosaka, Y. 2010. Metadata quality control in digital repositories and collections: criteria, semantics, and mechanisms. *Cataloging & classification quarterly*, 48(8), pp. 696-715.

Pascoal, A. P., Carvalho, R. B. and Xavier, E. M. A. 2013. Materialização do Perfil de Metadados Geoespaciais do Brasil em esquema XML derivado da ISO 19139. In: *XVI SBSR (XVI Simpósio Brasileiro de Sensoriamento Remoto*), Foz do Iguaçu - PR, Brazil, pp. 2441- 2448.

Rousidis, D. et al. 2014. Metadata for Big Data: a preliminary investigation of metadata quality issues in research data repositories. *Information Services & Use*, *34*(3-4), pp. 279-286.

Schindler, U. and Diepenbroek, M. 2008. Generic XML-based framework for metadata portals. *Computers & Geosciences*, 34(12), pp. 1947-1955.

Smits, P. C. and Friis-Christensen, A. 2007. Resource discovery in a European spatial data infrastructure. *IEEE Transactions on Knowledge and Data Engineering*, *19*(1), pp. 85-95.

Stvilia, B. et al. 2004. Metadata quality for federated collections. *Ninth International Conference on Information Quality.* Cambridge - MA, 2004.

Tani, A., Candela, L. and Castelli, D. 2013. Dealing with metadata quality: The legacy of digital library efforts. *Information Processing & Management*, 49(6), pp. 1194-1205.