

Large geospatial images discovery: metadata model and technological framework

Lukáš Brůha

Department of Applied Geoinformatics and Cartography,
Charles University in Prague
Albertov 6, 128 43 Prague 2, Czech Republic
lukas.bruha@natur.cuni.cz

Abstract

The advancements in geospatial web technology triggered efforts for disclosure of valuable resources of historical collections. This paper focuses on the role of spatial data infrastructures (SDI) in such efforts. The work describes the interplay between SDI technologies and potential use cases in libraries such as cartographic heritage. The metadata model is introduced to link up the sources from these two distinct fields. To enhance the data search capabilities, the work focuses on the representation of the content-based metadata of raster images, which is the crucial prerequisite to target the search in a more effective way. The architecture of the prototype system for automatic raster data processing, storage, analysis and distribution is introduced. The architecture responds to the characteristics of input datasets, namely to the continuous flow of very large raster data and related metadata. Proposed solutions are illustrated by the case study of cartometric analysis of digitised early maps and related metadata encoding.

Keywords: Catalogue systems, SDI, image data, metadata, automation.

1. Introduction

From the perspective of spatial data infrastructures (SDI) development strategies, the old maps represent a specific data source, especially regarding metadata. Cartographic documents always represented a unique way of expression and distribution of spatial information. In addition to their aesthetic value, old maps also yield significant scientific value. In these historical documents the methods for the representation of space at the time of their origin can be revealed. Moreover, these records provide valuable historical spatial data and document the temporal changes of represented phenomena, and so information about land use changes, the development of towns or river network transformations can be discovered. Such pieces of information are in the center of interest of present-day geospatial information systems. The International Cartographic Association (ICA) has even established the Commission on Digital Technologies in Cartographic Heritage to promote the collections of early maps to a general public. The group focuses on the development of methodologies applied to archiving, accessibility and analysis of thematic and geometric content of old maps.

Traditionally, the libraries have been the institutions, which were responsible for custodianship and the availability of map products through organized collections in their conventional form

(e.g. printed maps, journals, atlases, globes etc.). With the onset of digital era also the development strategies of libraries moved towards online distribution and constant request so as to provide improvable and innovative services for their users [27, 24]. Plenty of initiative has published early maps through geoportals, i.e. the David Rumsey Map Collection, the New York Public Library's instance of open-source Map Warper, the Geological Survey's Historical Topographic Map Explorer, the Office of Coast Survey's Historical Map & Chart Collection or the Alexandria Digital Library. Recently, some activities evolved aiming at the integration of digital map libraries with contemporary SDIs. Fernandez [10] presented an integrated access to the SDI through the crosswalk between geographic and bibliographic metadata profiles.

These issues regarding metadata interoperability and automation of data and metadata processing are also in the middle of interest of SDI solutions developers. The increase of spatial raster data volumes within SDI systems unveiled, how far the methods and tools originally designed for vector data are unsuitable for raster images, which are much bigger in data volumes and variable in storage formats [5, 22, 11]. The continuous flow of very large raster data from the digitization procedures of printed maps or satellite images receiving stations can be shown as an example of such data sources. These issues highly increase computational demands during any analytical task that requires additional data transfers, i.e. from the data store to the processing application. This implicates higher demands on the data store mechanism, cost-effective spatial images management and analysis approach. Also the availability of metadata, which supports support decisions on a dataset's appropriateness, alleviates the burden of users' requests on the system, allowing searching and assessing the data more effectively. The development of SDI is an active area of research regarding creation, update and authoring of such metadata [2, 26]. However, the development of the automation of metadata processing is still at the beginning and is being explored by many researchers [16, 18, 2, 14, 17, 21].

Accordingly, the solution presented here addresses the variability of metadata formats and differences between the standards used for data descriptions within libraries and geographical information systems (GIS). The model for metadata integration is introduced to link up the sources from these two distinct fields. This work describes the interplay between SDI technologies and potential use cases in digital libraries such as cartographic heritage.

Furthermore, this work proposes a technological solution focused on reduction of efforts related to the management of a continuous flow of large raster data and the creation of associated geospatial metadata. The architecture of the prototype system for automatic raster data processing, storage and distribution is introduced. This architecture responds to the characteristics of input datasets and to the demands for more effective spatial images management and analysis approach. Its characteristic feature is the shift of the application logic to the data store.

This article first discusses the metadata sources and proposes the metadata model for the integration of early maps metadata records within the SDI system. Further the technical and software requirements of SDI component for raster data are defined and the architecture, which is the basis for the system prototype implementation, is presented. Finally, the experiments and conclusions are reviewed.

2. Metadata model

Different standards with different focus are employed by librarian and SDI metadata catalogs. Currently, FGDC or ISO level standards are considered essential to spatial data infrastructures in order to satisfy the needs of archival, preservation and quality measure use cases. They are however, considered intricate and lengthy, without focus on discovery contexts such as web search or content description [12]. The insufficient metadata model for representation of the analysis outcome and the description of raster image properties can hinder reusability of existing analytical results and obstruct search for the desired image. In librarianship, the fundamental source of old map's description is represented by the bibliographic record. This description subsequently provides metadata for the digital environment usually encoded according to the MARC21 format. Such metadata is created in accordance with International Standard Bibliographic Description for Cartographic Materials (ISBD(CM)), Anglo-American Cataloging Rules with respect to local interpretation, i.e. AACR2R [1], or newer RDA [15] cataloging standard. Moreover, various cataloging procedures of old map collection [8, 19] and the final list of encoded metadata attributes are always depending on local specifics.

The survey of various metadata sources must be completed prior to the final constitution of the metadata model and the technological solution. Relevant information can be retrieved from multiple sources including raster headers, auxiliary files and external documents holding descriptive or technical metadata in a structured form.

2.1. Digital library and information science community metadata sources

Two kinds of metadata can be distinguished. First, the technical metadata describing the administrative and technical parameters of the digitization. Second, the actual bibliographic record entered mostly by the human description of the original document.

Technical metadata. The text documents of a well known format usually stored in an XML document are the source of technical metadata. These are usually produced in such a standardized format like Metadata for Images in XML schema (MIX) together with the image, during the extensive scanning campaign of archival materials. The owner of the original document, unique object identifier, data format, resolution, data volume and type of compression, identification and description of technical parameters of the scanner or the software used for scanning and post-processing belong among parameters encoded in such technical metadata documents. Metadata Encoding and Transmission Standard (METS) is a format frequently utilized by digital libraries to encapsulate logical connections between technical and bibliographical metadata.

Bibliographic records. The manual creation of some metadata like cataloguing procedure of early map collections cannot be avoided and represents an important source of metadata. Operators create metadata by writing descriptions of resources in a structured form, which can be automatically transformed afterwards. The bibliographical records are an essential resource for old maps description. During the cataloging procedure a plenty of descriptive attributes are recorded, including system identifiers, institution, authorship identification, title, physical description or scale, if available, and georeferencing entries etc. During the design of metadata model for discovery of geospatial resources, only a selection of all attributes were identified as relevant for such a purpose. The model focuses especially on elements describing the

mathematical foundation of cartographic materials like the scale, projection or geographical extent, which are encoded in 034 and 255 fields of MARC21 format. An example of a map scale encoding in MARC21 including the conversion of obsolete measuring units [19]:

```
255 $a Měřítka [ca 1:49 000]. 15,9 cm = 3 wiener Zoll = 1 oster. Meile v. 4000 Klafter od. 10000 Schr.  
255 $a Měřítka [ca 1 200 000]. 18,2 cm = 5 Meilen oder 20000 Klafter  
255 $a Měřítka [ca 1 1 126 000]. 4,6 cm = 7 militaria germanica communia  
034 $b 200000
```

For the sake of effective search for a desired digitized map, in addition to fields like author or period of creation, the content of the map and the time extent, which it is related to, the way of hypsography depiction or the language of the map are essential for cartographic document description. Through discussions with historical cartographers and geographers transpired the key role of proper representation of map series and map nomenclature, MARC21 fields 490 and 830. For instance The Third Military Survey of the Habsburg Empire, in which case the nomenclature is well established for searching the proper map sheets, moreover, it helps to avoid the language ambiguity between map sheets due to the change of map language during the survey. An example of the printed special map 1:75000 description [19]:

```
1101 $a Rakousko-Uhersko. $b Militargeographisches Institut $7 kn20060405003  
24510 $a Josefstadt und Nachod. $p Zone 4 Col. XIV $h [kartografický dokument] / $c K. u. k. militar-  
geographisches Institut ; Geripp. Scherling ; Terrain Neumann  
4901$a [Die Franzisco-Josephinische Landesaufnahme] 1 75 000  
830 0 $a Třetí vojenské mapování 1 75 000 ; $v 4XIV, 1903.
```

2.2. Metadata model for early maps discovery

Proposed final crossing between metadata elements of the digital library sources and ISO 19115:2003, see Table 1, respects the methodology of digitization and cataloging [19] of an old map collection of Faculty of Sciences, Charles University in Prague, which provided the source of bibliographic metadata. Further, it respects the minimal requirements of ISO 19115:2003 standard and INSPIRE directive for a valid record and it is compliant with the standard for digitization of cartographic documents of National Digital Library of the Czech Republic [25].

The elements of the resulting metadata record of an early map can be divided into three classes, as is indicated in Table 1. First, the bibliographic records originally created by cataloging operators. The technical descriptions encapsulated in the METS document bearing information about the map's scan quality, resolution, compression and scanning conditions, as well as an object identifier to facilitate linking the records back to the digital library system are the second source. Among technical metadata elements, the parameters generated by the system itself, i.e. data source URL, are assigned, including some preset values common for the whole dataset of early maps (*Contact*, *Role* or *Metadata language*). Beyond these two groups of metadata fields, which originate from the automatic processing line of early maps scans and related metadata, the product of automated or manual map analysis can be encoded within the *Reference system* and *Reference scale* fields or within *Supplemental field* for other results.

Table 1: The model for creation of an early map metadata document.

Fields of ISO 19115:2003	Source of values - fields of MARC21 document
Title	830 \$a\$v + 245 \$a
Date	008 (character position 07-10, 11-14)
Abstract	If available: 520 \$a Else: 245\$b\$c + 246\$a + 300\$a + 490\$a + 500 \$a
Keyword	043 \$a + 600 \$a + 610 \$a
Language	008 (character position 35-37)
Topic category	650 \$a + 651 \$a
Temporal extent	648 1 \$a (character position 1-4 a 5-9)
Geographic bounding box	034 \$d, \$e, \$f, \$g
	Source of values - technical metadata
Issue identification	METS document - field <ObjectIdentifier>
Identifier	UUID generated by metadata catalogue
Metadata language	"Cze"
URL	URL generated by map server
Metadata date stamp	System time of the document publishing
Organization name	Responsible party
Contact	Contact on responsible party
Role	Role of the responsible party
Hierarchy Level	"Dataset"
Statement (Lineage)	METS document fields <BasicImageInformation> + <ImageCaptureMetadata>
	Source of values - map analysis
Reference system	If available in MARC21 document: 255 \$b Else: Cartometric analysis output - EPSG code
Reference scale	If available in MARC21 document: 034 \$b Else: Cartometric analysis output - scale denominator
Supplemental information	Parameters of projection, scale and rotation If available: positional displacement, symbology type

3. Technical requirements for the implementation of SDI component for raster data

The automation of raster data processing, archiving, analysis and/or distribution requires the appropriate technological means. Following technical components were identified as essential for the design and implementation of the SDI module for raster image management:

- database system
- catalogue system
- visualization system
- administration system.

These components will be described in detail in following sections.

3.1. Database system

Within the database system the data model, which integrates both the data and metadata, is formed. Spatial database management system (SDBMS) capable of non-spatial data and georeferenced raster data storage is required to accommodate such a data model. Database tables would store raster datasets in a native format of the selected SDBMS platform, metadata tables would hold the ISO 19115:2003 elements. The current usual practice utilizes the out-of-the-database raster data storage, when only related metadata is stored in a relational database, while rasters are kept in the original raster binary format. The in-database strategy [20, 29] employed by the proposed solution stores images in a native database format, thus moving the image processing closer to the data and allowing for both concurrent, and parallel data processing. Database platforms with in-database raster data storage support also natively provide computationally optimized functions for raster data manipulation, editing and yielding of image statistics or image histogram. This functionality is re-usable within the data store and can be incorporated into the raster-based analytical procedures to be developed, making their implementation straightforward and their computational performance efficient. The functional requirements on SDBMS platform are summarized as follows:

- spatial indexing
- raster bands accessors
- raster pixel accessors and setters
- raster band statistics
- datum definition and coordinate system transformation
- storage of attribute data and metadata to the stored spatial data
- support for Geospatial Data Abstraction Library (GDAL) for raster format load and export operations
- raster pyramids and tiling support
- extrusion of raster regions
- storage of geo-referenced rasters and vector data
- analytical tool for detection of intersections with vector data
- map algebra over individual pixels.

The chosen database platform is also supposed to provide tools for communication with desktop and web applications. PostgreSQL with spatial extension PostGIS was chosen for the implementation of the proposed SDI concept. PostGIS raster extension supports raster data manipulation and complies with the stated requirements. This open-source solution is widely used and involves stable and active community. It ensures a future development and also easy cooperation with other academic institutions.

3.2. Catalogue system

Catalogue servers or metadata solutions, that include server side, are the second component required for the system design. The common objective for metadata catalogue usage within the SDI is to integrate all metadata records of disparate data sources and distribute them in a compact manner. The requirements laid on the catalogue are:

- the support of management and administration of the metadata
- the ready to use web graphic user interface (GUI) to interact with the end users
- extensibility of metadata profile templates
- the searching mechanism supporting the filtration based on multiple parameters
- metadata import via eXtensible Markup Language (XML) services.

GeoNetwork opensource catalogue is the commonly used catalogue solution. It fits to the requirements stated above and was finally selected for the integration with the prototype system proposed in this work.

3.3. Visualization system

The third key component of the proposed solution is the server technology capable of map distribution. This mapping application would be connected with the spatial database and would play the role of middleware by delivering the data to the client-side, providing on-line mapping service for the SDI. The requirements laid on the map server include:

- web administration environment for map layer management;
- Representational State Transfer Application Programming Interface (REST API) to programmatically manage the server;
- OGC compliant Web Map Service (WMS) support
- role-based access control to authorize users or groups of users.

Common use of the GeoNetwork metadata catalogue is in conjunction with the GeoServer map server, which act as a source of published georeferenced images (WMS) or complex features (Transaction Web Feature Service). Well documented and easy to deploy integrated system of GeoNetwork and GeoServer together with the fulfilment of the requirements stated above are behind the decision to employ the GeoServer as a map server application for the prototype solution.

3.4. Administration system

The design and implementation of an executable application would be another prerequisite to automation of the continuous flow of raster data and related metadata processing. Its objective would be to initialize and control the manipulation of raster data and provide an integrated approach to metadata generation. Operating system independence is the major requirement laid on such an application.

3.5. Architecture of the prototype solution

The architecture of the automatic raster data and metadata management system is depicted in Figure 1. This architecture is composed of three main layers:

- administration layer
- storage layer
- service layer.

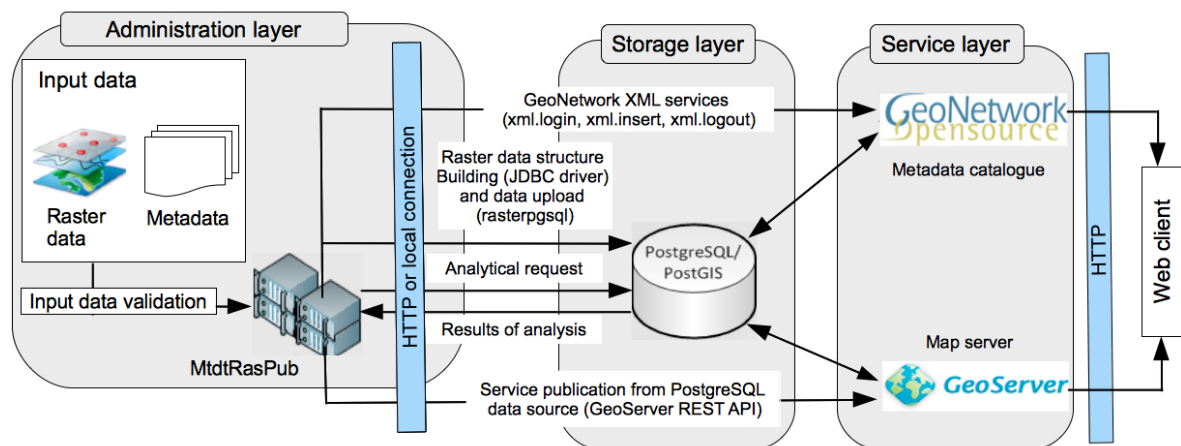


Figure 1: Architecture of the raster-based SDI module prototype system.

Administration layer

Administration layer provides an environment for initialization and configuration of the complete solution for automatic raster data and metadata processing and publishing. Technically, the administration layer is based on Java application (MtdtRasPub). The application continually checks for unprocessed data and triggers the building of database structures, the import of data and metadata to the database and the web map service publishing, when such data are detected. The administration layer is also responsible for management of related metadata. Procedures deployed within this layer take advantage of known structure of the metadata source documents and use xPath expressions to harvest the required elements defined in metadata model. The administering application executes requests for analysis of raster data. Such requests are sent to the data store, where the analytical logic is deployed. The in-database concept allows avoiding moving large data sets from the databases to detached analytical software. The results of analytical procedures or the metadata describing their output are received back.

Storage layer

The storage layer contains the databases, which store the spatial data and metadata. The PostgreSQL database platform is employed by the prototype solution.

Metadata database. According to the ISO 19115:2003 specification, the metadata records are stored in the XML structure. For the sake of automation of storage, update and pub-

lication of metadata together with the corresponding data, metadata needs to be stored in relational database. The metadata document is formed within the MtdtRasPub application and consequently GeoNetwork's XML services (xml.metadata.insert operation) are utilized for update or creation of new records in the catalogue.

Geospatial raster database. The raster data storage model utilizes the native PostGIS Raster format to support the data analysis provided by the database platform. The communication between the database and the map server is achieved by the use of GeoTools Image Mosaic JDBC plugin (IMJDBC), which requires the defining geometry for every raster tile placement. PostGIS Raster function `ST_Envelope` is applied to create the geometry. This way of storage allows for manipulation with all PostGIS raster functions and simultaneously the publication through GeoServer map server.

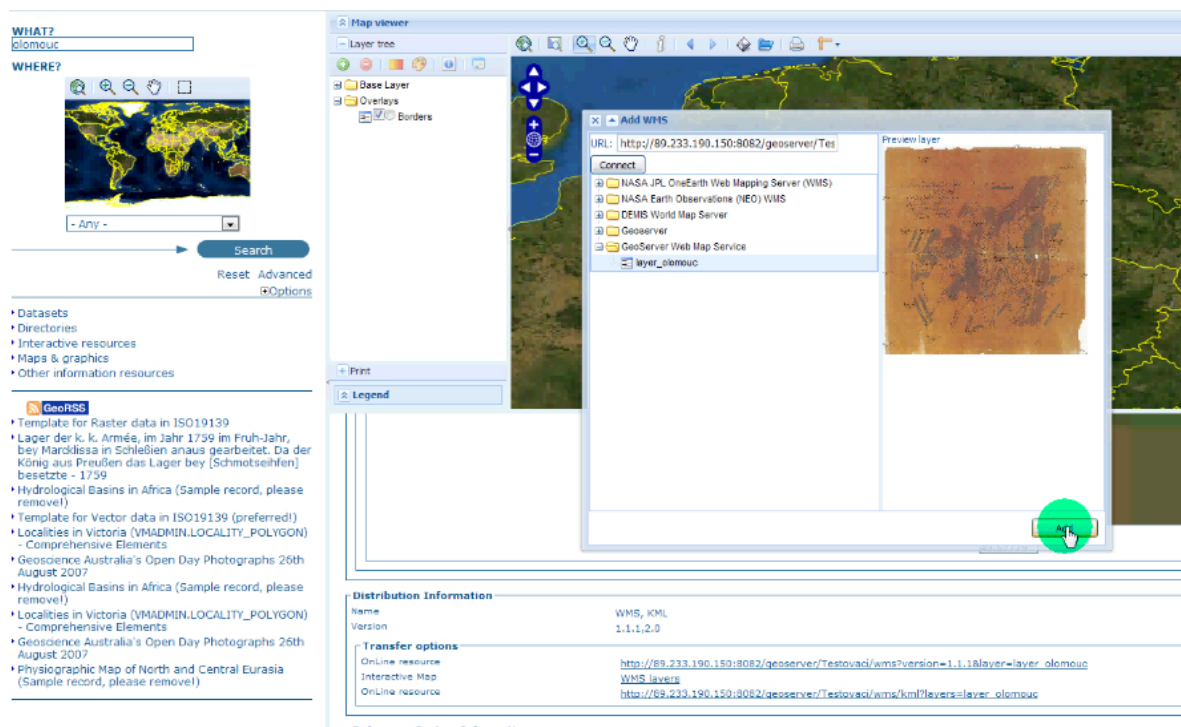


Figure 2: The overall view on the GeoNetwork's GUI with integrated metadata search, the overview of selected map's metadata and map visualization window.

Service layer

The objective of this service layer is to discover and publish metadata records and raster sources from PostgreSQL database through OGC compliant Web Services.

Metadata publishing. The deployed GeoNetwork opensource provides OGC Catalogue Services for the Web (CSW) server to access the metadata within PostgreSQL and enable search and update operations. This server supports 2.0.2 version of the OGC specification, supporting GetCapabilities, DescribeRecord, GetRecordById, GetRecords, Harvest and Transaction CSW operations.

Map service publishing. GeoServer supports as data source the rasters of three following data types - geotiff, worldimage and imagemosaic. To accommodate the needs of automatic publication of the data source retrieved from PostGIS Raster, the extension of supported data formats is necessary. The opensource nature of GeoServer allows adding among the supported formats the imagemosaicjdbc applying IMJDBC GeoTools plugin.

Graphic user interface

The architecture of the raster-based enhancement of SDI described above also provides a wide range of possible GUIs for user interaction with the prototype solution. GeoServer and GeoNetwork are modular open-source components following OGC standards and providing means for integration with desktop and web-based GIS systems, including web-based user interfaces. The GUI of GeoNetwork opensource, depicted in Figure 2, supports interactive user editing of ISO 19115:2003 metadata elements providing automatic updates to the database through CSW service. Integrated map window allows user to visualize the looked-up map and even combine it with another data available. GeoServer's GUI enables organization of individual map layers into group of layers for a signed-in user. These environments provide effective tools for constitution of map compositions from several layers as for the map purpose and user's needs.

Table 2: Metadata record of *Školní nástěnná mapa Afriky F. Umlaufa a J.G. Rothauga*.

Title	Školní nástěnná mapa Afriky F. Umlaufa a J.G. Rothauga
Date	1900
Abstract	pro české střední školy, upravil Josef Krejčí, 1 mapa, legenda
Keyword	regionální geografie; politická geografie; správní obvody; didaktika geografie; Afrika; nástěnné mapy; školní mapy
Language	Cze
Topic category	nástěnné mapy; školní mapy; Afrika
Temporal extent	1900
Geographic bounding box	-26.366944; 63.808333; -38.548056; 45.583056
Issue identification	1165964
Identifier	237
Metadata language	Cze
URL	http://mapy.natur.cuni.cz:8080/geonetwork/srv...
Metadata date stamp	2015-07-29T15:29:44
Organization name	CZ PrCU
Contact	mapcol@natur.cuni.cz
Role	point of contact
Hierarchy Level	Dataset
Statement (Lineage)	5250x4660; sRGB; IEC61966-2.1 2012-12-06; ScannTech 401i

4. Discussion and system evaluation

A solution for automatic management of continuous flow of large raster data and related metadata has been proposed. The presented technological framework allows to integrate

descriptive information about images from various sources based on the metadata model and produces standardized geospatial metadata records. As this work focused on early maps, the metadata model to represent the corresponding map's description was introduced based on recommendations from librarians and cataloging operators. Table 2 shows a metadata record in ISO 19115:2003 standard of one particular dataset.

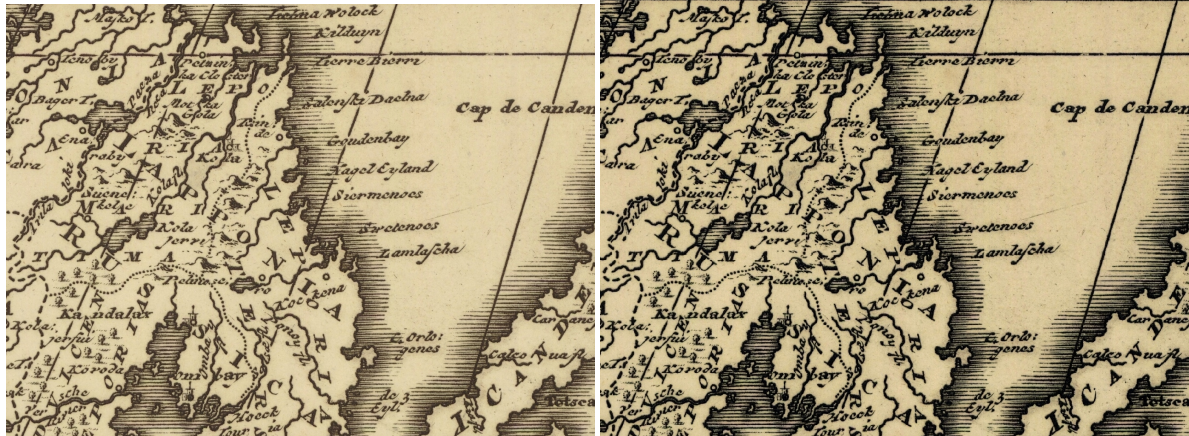


Figure 3: The segment of an early map (a) before and (b) after application of image histogram stretching procedure.

The technological solution of the prototype system is based solely on opensource technologies. While this approach has several advantages compared to proprietary systems, including the opportunities for customization and a support of active communities of developers, it also causes a number of technological challenges. Due to the heterogeneity of independently developed components, which were not implemented with mutual respect, the customization of the system components or additional libraries are required to enable the communication between all layers. That complicates the final configuration of the server solution. Another potential complication, identified also in works concerned with representation of metadata related to vector data [21], is the GeoNetwork's storage mechanism of metadata. The whole XML string is stored in a single column, thus accessing a single metadata element can only be done by GeoNetwork's parsing mechanism.

The consideration, whether the out-of-the-database or in-database raster data storage should be employed, must consider the major objective of such application. If the solution is meant for mere distribution of raster data, the out-of-the-database storage provides slightly better time responses and poses less complications when building the server solution and communication channel between GeoServer and PostgreSQL/PostGIS. Because the leading motivation was such a system, that would support the raster images analysis providing native database functions for raster data manipulation, editing and image statistics, the in-database approach to the data storage was chosen. This facilitated the development of new analytical procedures within the data store.

Image enhancement. Some tests of the in-database approach, like an image histogram equalization or cloud and snow detection from satellite imagery procedures, were presented in previous author's work [5]. The histogram equalization can be applied to improve the legibility of an old map or to remove differences in contrast between separate map sheets of map series

for the sake of mosaicking, see Figure 3 for an example. Consequently, in the metadata field *Supplemental information* new intensity values of the map sheet can be retained.

As the approach focuses on digitized early maps and related metadata, the character of early map (fragility of the original and its uniqueness regarding the means of map representation or map construction) however suggests the human mediation in the creation of descriptions of some raster sources cannot be completely avoided. Nevertheless, the proposed metadata model facilitated to represent analytical results obtained manually or with a use of an external application.

Cartometric analysis. The cartometric analysis of an old map is another use-case, whose results are desired to be recorded for the future discovery. Cartometric analysis is a key prerequisite for any proper analysis of a map content. Many authors, [7, 6, 28] to name a few, paid attention to development of cartometric analysis methods, so as to allow for environmental research, land use changes description and many other discoveries based on old maps. This experiment was focused on the assessment of planimetric accuracy and the estimation of cartographic parameters of old maps. Detected projection, category of projection, geographical coordinates of the cartographic pole, the latitude of the true parallel, the longitude of the prime meridian or percentage of points fitting used geometrical model are among parameters examined by the procedure. Theoretical background to the applied methodology is available in publications [23, 4]. The cartometric analysis procedure was performed by external software package *detectproj* [3]. Other parameters of map analysis, i.e. map symbology description, completeness of map content or positional accuracy can be done manually [13]. The deviation of the graticule of Delisle map [9] and the graticule of estimated stereographic projection is depicted by Figure 4.

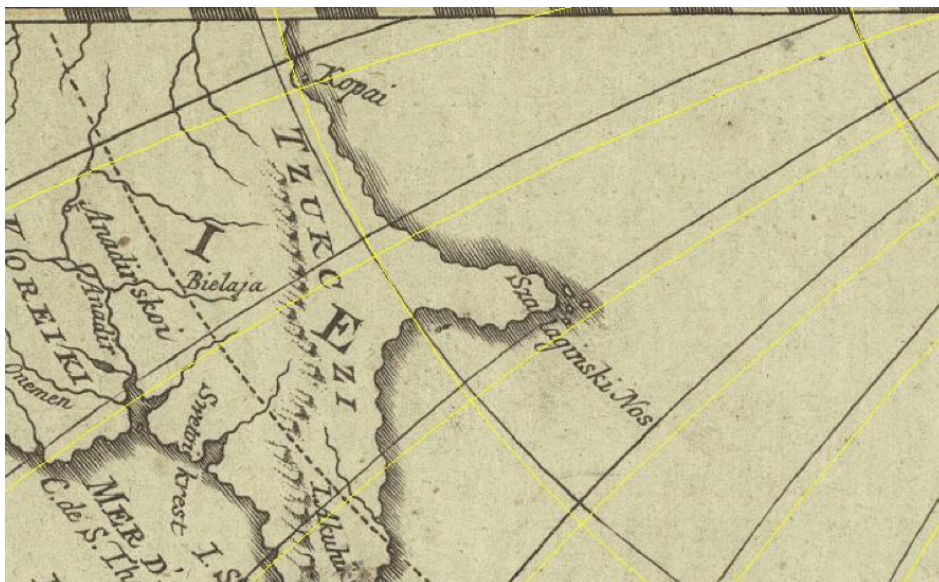


Figure 4: The visualization of the deviation between the graticule of Delisle (1774) map and the graticule of estimated stereographic projection. Figure drawn by [13].

Another example illustrates the measurement of positional displacements of selected cities, which is depicted in Figure 5.

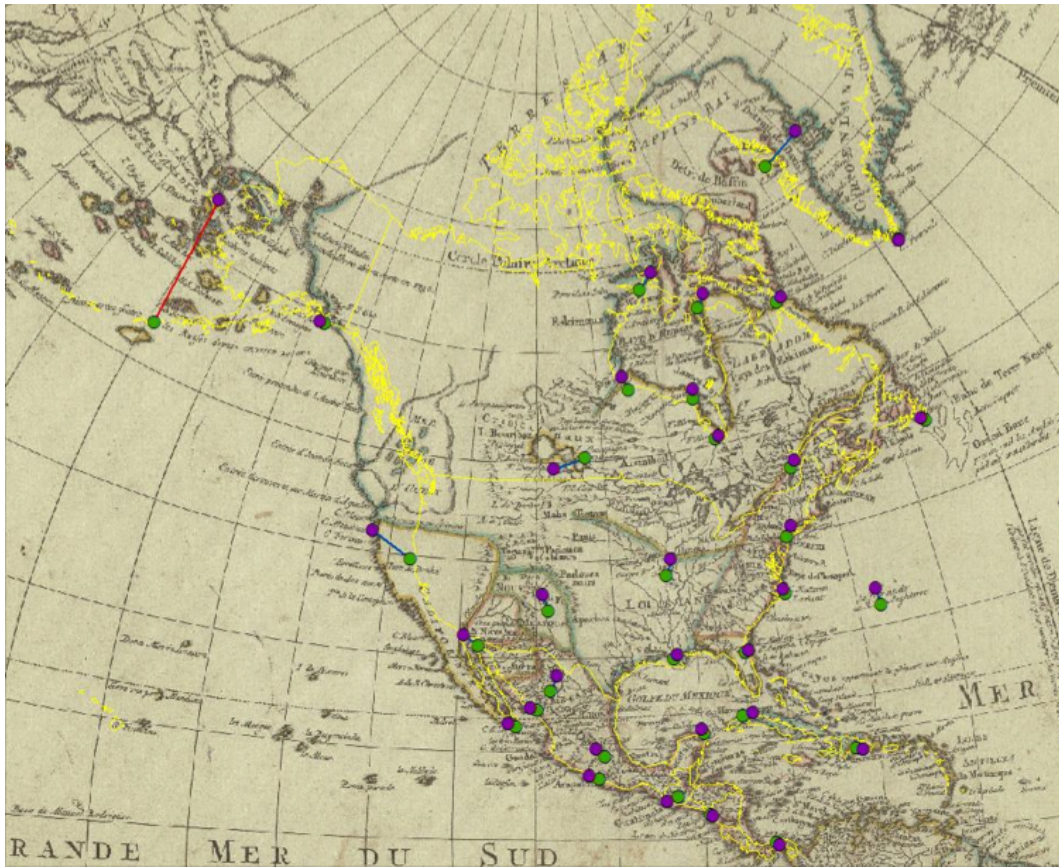


Figure 5: The illustration of the shift of points and the continent outline in Delisle map in comparison to the actual state. Figure drawn by [13].

The content of the *Supplemental information* field related to the Delisle map analysis and its visualization by the GeoNetwork metadata catalogue is demonstrated in Figure 6. The most probable detected projection, latitude of the true parallel, the longitude of the prime meridian, geographical coordinates of the cartographic pole, HOMT - standard deviation on identical points after homothetic transformation, HELT - standard deviation on identical points after Helmert transformation and measured positional displacements on selected places were encoded.

5. Conclusion

Efforts for disclosure of map collections are one of the very actual aspects of the SDI use. Moreover, it is also an impulse for further development of metadata interoperability and tools for effective processing of continuous data-flows of raster datasets generally. These efforts were presented through the example of the map collection of the Charles University in Prague and the SDI of the Faculty of Science. The interplay between SDI technologies and librarianship was described in this work and presented on some examples of a cartographic heritage discovery.

Introduced solution met the needs for automation of a continuous flow of very large raster data

Projection & Scale analysis:
1 stere 0.0 -89.5 0.0 0.0:
Scale HOMT: 42865490.7
Scale HELT: 42865587.5
Rotation HELT: -0.12 deg

Positional displacement [km]:
Ciudad de Panamá 48,91; Cape Hatteras 58,29; León 26,76;
Bermuda 186,23; Ciudad de Guatemala 153,30;
New York City 119,18; Acapulco 130,65; Québec 85,50;
Ciudad de México 145,64; Cape Race 45,80; Cabo Catoche 53,48;
Dauphin 357,84; La Habana 117,50; York Factory 150,40;
Santo Domingo 76,14; St. Augustine 45,90; New Orleans 54,29;
Unalaska Island 1130,73; Santa Fe 202,70; Cape Mendocino 532,35

Figure 6: The content of *Supplemental information* field within GeoNetwork catalogue.

and associated metadata. And, as a result, it saved custodians' time and limited the space for manually introduced errors. The released human resources can concentrate on descriptive metadata creation to support semantic search and enhance the raster data discovery. The actual efforts aim at server-based deployment of the solution. The future development founded on proposed technology and architecture will be focused at the extension of content-based image analyses followed by automated content-based metadata creation.

Acknowledgement

This work has been supported by the TEMAP project of the NAKI program of the Ministry of Culture of the Czech Republic with DF11P01OVV003 code.

References

- [1] Ivana Andresová. *Anglo-americká katalogizační pravidla: příručka pro katalogizátora s příklady ve formátu UNIMARC a MARC 21*. Praha: Národní knihovna české republiky, 1994. ISBN: 80-705-0187-1.
- [2] James K. Batcheller. "Automating geospatial metadata generation—An integrated data management and documentation approach". In: *Computers & Geosciences* 34.4 (Apr. 2008), pp. 387–398. DOI: [10.1016/j.cageo.2007.04.001](https://doi.org/10.1016/j.cageo.2007.04.001).
- [3] Tomáš Bayer. *Detectproj: estimation of the cartographic projection and its parameters from a map*. Software. 2015.
- [4] Tomáš Bayer. "Estimation of an unknown cartographic projection and its parameters from the map". In: *GeoInformatica* 18.3 (Mar. 2014), pp. 621–669. DOI: [10.1007/s10707-013-0200-4](https://doi.org/10.1007/s10707-013-0200-4).
- [5] Lukáš Brůha. "Automation of geospatial raster data analysis and metadata updating: an in-database approach". In: *GEOGRAPHICA* 49.2 (Nov. 2014), pp. 49–56. DOI: [10.14712/23361980.2014.14](https://doi.org/10.14712/23361980.2014.14).

- [6] Jiří Cajthaml and Jiří Krejčí. “Využití starých map pro výzkum krajiny”. In: *Sborník symposia GIS Ostrava 2008*. Ostrava: VŠB-TU Ostrava, 2008.
- [7] Jiří Cajthaml et al. “Georeferencing and Cartographic Analysis of Historical Military Mappings of Bohemia, Moravia and Silesia”. In: *CTU Reports, Proceedings of Workshop 2007*. Praha: Czech Technical University, 2007.
- [8] Joan Capdevila et al. “Gateway MARC21-ISO19115: definition and reference implementation”. In: *e-Perimtron* 7.3 (2012), pp. 155–162.
- [9] Guillaume Delisle. *Carte d’Amerique: divisees en ses principales parties*. Amsterdam: Covens & Mortier & Covens junior. 1774.
- [10] Alberto Fernández-Wytenbach et al. “First approaches to the usability of digital map libraries.” In: *e-Perimtron* 3.2 (2008), pp. 63–76.
- [11] Stanislav Grill et al. “Archive and catalogue system for receiving satellite data as a part of academic SDI”. In: *Imagin[e,g] Europe*. Ed. by Manakos Ioannis and Kalaitzidis Chariton. 2010.
- [12] Darren Hardy and Kim Durante. “A Metadata Schema for Geospatial Resource Discovery Use Cases”. In: *Code4Lib* 25 (2014).
- [13] Petra Jílková. “Severní Amerika ve starých mapách”. Bachelor Thesis. Charles University in Prague, Czech Republic, 2014.
- [14] Mohsen Kalantari, Abbas Rajabifard, and Hamed Olfat. “Spatial metadata automation: a new approach”. In: *Spatial science conference*. Ed. by et al. B. Ostendorf. Adelaide, Australia: Surveying & Spatial Sciences Institute, 2009, pp. 629–635.
- [15] Judith A. Kuhagen. *RDA essentials*. 2010.
- [16] Miguel Manso et al. “Automatic metadata extraction from geographic information”. In: *AGILE 2004: 7th conference on geographic information science: conference proceedings*. Ed. by Toppen F. Heraklion: Crete Univ. Press, 2004.
- [17] Miguel Manso-Callejo, Mónica Wachowicz, and Miguel Bernabé-Poveda. “The Design of an Automated Workflow for Metadata Generation”. In: *Metadata and Semantic Research*. Springer Science + Business Media, 2010, pp. 275–287. DOI: [10.1007/978-3-642-16552-8_25](https://doi.org/10.1007/978-3-642-16552-8_25).
- [18] J. Nogueras-Iso et al. “Metadata standard interoperability: application in the geographic information domain”. In: *Computers, Environment and Urban Systems* 28.6 (Nov. 2004), pp. 611–634. DOI: [10.1016/j.compenvurbsys.2003.12.004](https://doi.org/10.1016/j.compenvurbsys.2003.12.004).
- [19] Eva Novotná. “Staré mapy a grafiky v Geografické bibliografii ČR on-line”. In: *Knihovna : knihovnická revue* 24.1 (2013), pp. 5–27.
- [20] O. R. Obe and L. Hsu. *PostGIS 2.0 3D and Raster support enhancements*. (Visited on 08/17/2015).
- [21] Hamed Olfat et al. “A GML-based approach to automate spatial metadata updating”. In: *International Journal of Geographical Information Science* 27.2 (Feb. 2013), pp. 231–250. DOI: [10.1080/13658816.2012.678853](https://doi.org/10.1080/13658816.2012.678853).

- [22] Du A. Peijun and Chen B. Yunhao. *ISPRS Workshop on Service and Application of Spatial Data Infrastructure, XXXVI(4/W6), Oct.14-16, Hangzhou, China. Some key techniques on updating spatial data infrastructure by satellite remote sensing imagery*. 2005.
- [23] Markéta Potůčková and Tomáš Bayer. “Application of e-learning in the TEMAP project”. In: *Geoinformatics FCE CTU 9* (Dec. 2012), pp. 91–99. DOI: [10.14311/gi.9.8](https://doi.org/10.14311/gi.9.8).
- [24] Muthu Sabarish Senthilnathan, Gkadolou Eleni, and Stefanakis Emmanuel. “Historical Map Collections on Geospatial Web”. In: *GEOMATICA 67.3* (2013), pp. 163–171. DOI: [10.5623/cig2013-035](https://doi.org/10.5623/cig2013-035).
- [25] Pavla Švástová and Jaroslav Kvasnica. *Definice metadatových formátů pro digitalizaci monografických dokumentů (monografií, kartografických dokumentů, hudebnin)*. Tech. rep. Praha: Národní knihovna, 2013.
- [26] Sergi Trilles et al. “Assisted generation and publication of geospatial metadata”. In: *Proc.15th AGILE Int. Conf. Geographic Information Science: bridging the geographic information sciences*. New York: Springer, 2012, pp. 24–27.
- [27] Ifigenia Vardakosta and Sarantos Kapidakis. “The New Trends for Librarians in Management of Geographic Information”. In: *Procedia - Social and Behavioral Sciences 73* (Feb. 2013), pp. 794–801. DOI: [10.1016/j.sbspro.2013.02.120](https://doi.org/10.1016/j.sbspro.2013.02.120).
- [28] Bohuslav Veverka and Lucie Šrajerová. “Kartometrická analýza polohopisné přesnosti geografického obsahu historické Komenského mapy Moravy”. In: *Sborník 18. kartografické konference*. Olomouc, 2008.
- [29] Qingyun Xie et al. “In-database image processing in Oracle Spatial GeoRaster”. In: *ASPRS 2013 Annual Conference*. Baltimore, Maryland., 2013.