

# A BROKERED VIRTUAL HUB APPROACH FOR THE GENERATION OF WEB APPLICATIONS BASED ON HISTORICAL MAPS

Mattia Previtali<sup>1</sup>, Miguel Ángel Latre<sup>2</sup>

<sup>1</sup> Politecnico di Milano, Department of Architecture, Built Environment and Construction Engineering, Via Ponzio 31, 20133 Milano, Italy – [mattia.previtali@polimi.it](mailto:mattia.previtali@polimi.it)

<sup>2</sup> Universidad de Zaragoza, Department of Computer Science and Systems Engineering, C/ María de Luna 1, Zaragoza, E-50018, Spain - [latre@unizar.es](mailto:latre@unizar.es)

## Abstract

Integration between historical maps and current cartography is nowadays recognized of primary importance in many applications (e.g., urban planning, landscape valorisation and preservation, land changes identification, etc.). However, due to large variety in Geographical Information (GI) standards and interfaces for data publishing, some technical issues arise for developers when integrating different data for the generation of new web based applications. In addition, information overload makes difficult their discovery and management: without knowing the specific repository where the data are stored, it is difficult to find the information required. To partially cope with those problems this paper describes a new brokering based approach for the generation of web applications based on multi temporal GI data gathered from different providers. In particular, this new approach is exemplified by a couple of new web applications built on top of the developed solution. Even if the two applications deal both with historical maps, they present significant differences in technical (e.g., libraries, development environment, data formats, etc.) and non-technical (e.g., user addressed, user requirements, etc.) aspects showing the flexibility of the solution.

## Keywords

Open data, brokering approach, interoperability, historical maps, web services

## Acknowledgements

The research leading to the results of this paper is partially funded under the ICT Policy Support Programme (ICT PSP) as part of the Competitiveness and Innovation Framework Programme by the European Community (CIP) GA n°620400.

## 1. Introduction

Geospatial data are more and more widespread. Nowadays, many different institutions, such as geographical institutes, public administrations, collaborative communities (e.g., OpenStreetMaps) and web companies, make available a large number of maps in many different formats (raster images, vector data, catalogues, textual data, etc.), see also Bonham-Carter, 2014. However, this heterogeneity of the geo-information world is also widely recognized as a significant barrier for a wide use of Geographical Information (GI). Indeed, different data formats, services and interfaces are used by different communities of users and the integration of data coming from various sources is in many cases a complex task (Goodchild et al., 2012). Heterogeneity in the GI world can be considered de-facto. Indeed, different user communities have different requirements in terms of communication protocols and technology, limiting a

1 fruitful integration of Open Data (OD). Despite this, in recent years many important programs and  
2 initiatives tried to address the ambitious objective of fostering standardization and interoperability.  
3 Examples of these initiatives are INSPIRE Directive established for geo-information of European Public  
4 Authorities, GMES/Copernicus core and downstream services for environmental and security applications  
5 based on European Earth Observation data, the Global Earth Observation System of Systems (GEOSS), Eye  
6 on Earth and Shared Environmental Information System (SEIS) initiatives of the European Environmental  
7 Agency (EEA). However, even if these initiatives produced in recent years several very good standards (e.g.  
8 from OGC, ISO, etc.), the interoperability issue between different user communities still exists. For this  
9 reason, the opportunity to access and integrate multiple sources (e.g. downstream services, open datasets  
10 from national/local Spatial Data Infrastructures (SDI), etc.) for the development of new applications,  
11 requires the need to address different standards and protocols, which may be a quite complex and  
12 cost/time consuming task (Guigoz, 2016).  
13  
14  
15

16  
17 An example of interaction between different GI domains is represented by the integration between ancient  
18 and historical maps with current cartography. Nowadays, importance of ancient and historical maps is not  
19 only connected to their intrinsic historic and artistic value but also as a source of information to derive the  
20 former organization of the territory (urbanized areas, field cultivations, etc.) as well as their evolution over  
21 time (Grosso, 2010). Since territory and urban areas are in continuous evolution, to understand on-going  
22 dynamics, the study of the current state is not sufficient for their protection, conservation and future  
23 planning (Cerreta and Mele, 2012). In order to understand the evolution over time of these complex  
24 structures, historical maps are an important tool and can be efficiently used as a reference for urban  
25 planning and territorial management as well as urban and landscape preservation and land changes  
26 identification. However, historical data are in many cases mainly published as bibliographic data-sources,  
27 without taking into consideration their intrinsic cartographic content and their integration with  
28 cartographic products for the production of web services is not straightforward. In addition, few catalogue  
29 services exist and in spite of a large variety and availability of historical data, information fragmentation  
30 makes the discovery of datasets a challenging task: without knowing the specific repository where the data  
31 are stored, it is difficult to find the information required. Furthermore, problems can also arise due to  
32 restricted interoperability between different data sources. For example, either Web Feature Service (WFS)  
33 or OpenSearch (an Atom-format with download link) are frequently used for downloading datasets  
34 (Technical Guidance for the implementation of INSPIRE Download Services) but their integration is not a  
35 trivial task.  
36  
37  
38  
39  
40  
41  
42

43 Another important aspect is given by the fact that many available web platforms and services dealing with  
44 historical maps generally allow for a simple visualisation of the maps and no further knowledge is extracted  
45 from them. On the other hand, Historical GIS (HGIS) proved to be one of the most promising frameworks,  
46 which can provide innovative tools and methodological approaches in this field, since they allow a  
47 complete and exhaustive framework of transformations occurred (Southall 2014).  
48  
49  
50

51 This paper presents two applications dealing with historical maps relying on a brokering based approach to  
52 access in a seamless way both historical maps (served by different data providers) as well as current  
53 cartography. These applications rely on the approach developed in the European project ENERIGIC OD  
54 (Competitiveness and Innovation Framework Programme by the European Community – CIP – GA  
55 n° 620400). To provide a single point of access to heterogeneous data sources, the developed approach  
56 uses a virtual node, a gateway based on a broker approach (Boldrini et al., 2015; Nativi et al., 2013), where  
57 users can seamlessly discover and access potentially unlimited datasets. The paper is organized as follows:  
58 Section 2 describes the workflow developed for creating web applications on the top of a geographic  
59  
60  
61  
62  
63  
64  
65

broker (Virtual Hub), Section 3 describes the main characteristics and components of the Virtual Hub used to develop the two applications. Section 4 presents a short literature review on existing brokering platforms of GI and discusses them in relation with the brokering framework used in the development of the applications presented in Section 5: (i) GeoPan (§ 5.2) and (ii) Zaragoza Histórica (§5.3), after a description and discussion on the data used in the applications and their processing (§5.1). Advantages and further discussions of building web applications on the top of a geographic broker are discussed in Section 6.

## 2. Overview of the adopted approach

This paper presents the development of two new applications on the top of a brokering-based component, called Virtual Hub (VH). Both applications use ancient cartography and historical maps in conjunction with current cartography. Those datasets are accessed through these VHs, that that allow to access seamlessly to datasets published by different providers.

A general trend in the development of programs that share OD is to rely on a Service Oriented Architecture (SOA). In this architecture, users play the role of service consumers, data producers represent service providers, while the cyber-infrastructure acts as a service clearinghouse or service registry, which is generally kept as simple as possible. In particular, the cyber-infrastructure consists of a set of few selected common international standards (i.e. Web services protocols), common data models, and rules (e.g. best practices) adopted by both data producers and users to publish available resources, discover them and bind them. A typical approach used for the development of web applications in this archetype is based on a stateless Client-Server (C-S) architecture. In this configuration, client components send specific requests to servers that offer data services; requests to different servers are performed by using one of the specific interfaces supported (Fielding, 2000). The interoperability issue is solved by working on defining these interfaces and boosting their adoption as standards at different levels (Schmidt et al., 2005; Ortiz, 2007). This approach to interoperability is generally defined as “Federated Architecture”. In fact, it is based on the definition of a common or federal data model and common/federal interoperability protocol(s). In the GI domain, most SDIs around the world and initiatives working on GI interoperability (e.g. INSPIRE) are based on this model. They use the SOA concept to define a limited set of standard specifications defining data model, data structure and interfaces to interconnect system components (i.e. clients and servers). In other words, federated systems address interface heterogeneity by pushing adoption of common standards.

The SOA approach represents a top-down solution to the interoperability issue and it works very well in consistent and controlled frameworks (e.g. enterprise environments), or where the approach is embedded in a strong legal framework that makes it mandatory for stakeholders to adopt the agreed standards and protocols (e.g. the NSDI for US federal government, and INSPIRE in Europe). However, this approach faces some issues when dealing with practical applications based on data integration provided by different communities, like the ones discussed in this paper. Indeed, each user community generally tries to develop its own infrastructures using its own standards according to its specific requirements (Fig. 1). Indeed, it is noteworthy that the lack of agreed interoperability standards is not due to removable barriers, such as attitude, legal, financial or technological barriers. It is inherently impossible to define a common standard which is suitable for all the possible systems handling geo-information. This is one of the reasons why, although several very useful standards have been developed in the recent years (e.g. from OGC, ISO, etc.), they are not adopted by all the different domains. As a consequence, the opportunity to access and integrate multiple sources for the development of new applications requires sometimes to work with data published according to different standards and protocols, which may be a quite complex and expensive task.

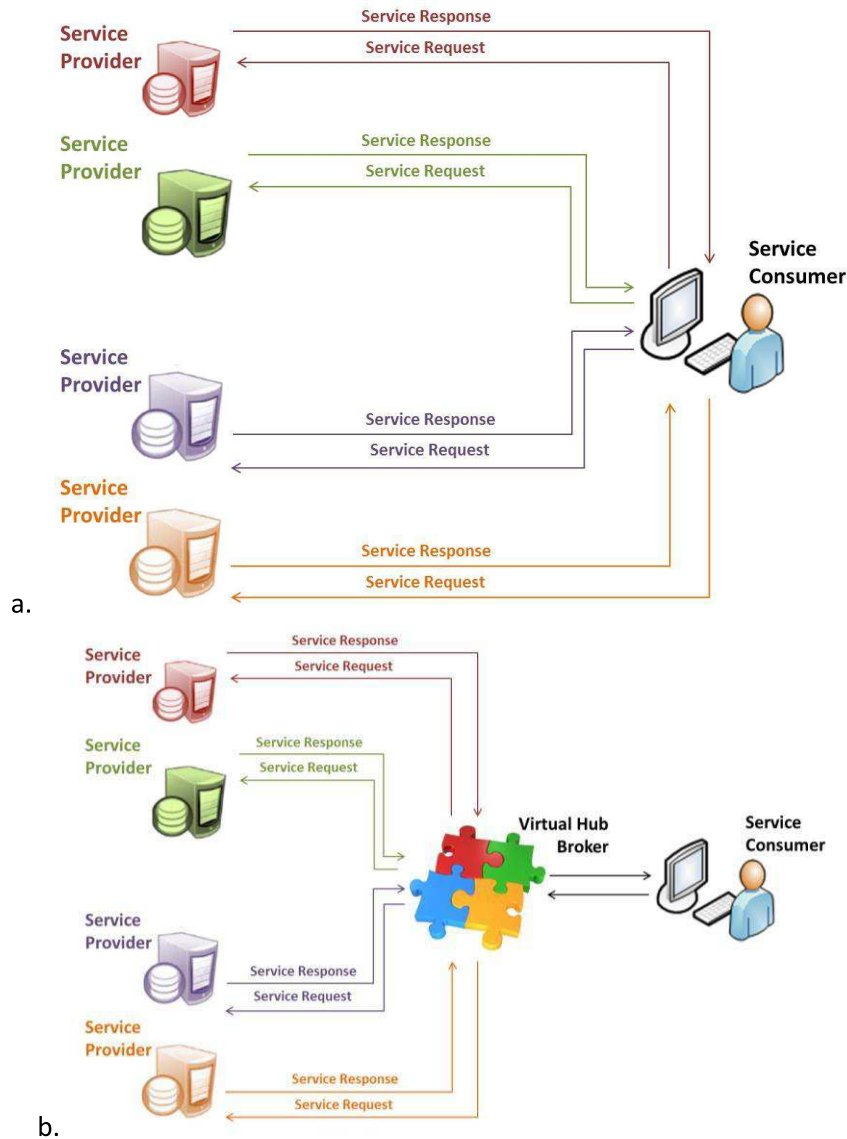


Fig. 1: The difference between a Service-Oriented-Architectures-SOA approach (a) and a brokering approach (b).

In order to remove this barrier, a bottom-up solution to interoperability is presented in this paper (Section 3) which is based on a brokered architecture developed by the European project ENERIGIC OD. This solution tried to partially cope with the previously listed problems by developing a mediation system based on a brokering framework. This mediation layer represent a virtual node, called Virtual Hub, used to allow discoverability and accessibility to datasets provided according to different standards and though different publishing protocols in a seamless way. The advantage of this approach for an application developer is that entry barriers (like the one represented by the need to implement access to data through standards that may be difficult and often unknown to the developer) are lowered, allowing thus for a wider use of OD. In addition, specific services are integrated that provide specific functionalities (e.g., semantic search, multilingual support, reprojection, etc.). Interaction with the Virtual Hub is performed by using a set of APIs (REST and/or Web) that are exposed by the system. Those API allow the development of new geographically-based applications. The potentiality of this brokering approach is exemplified by two applications dealing with historical maps built on top of the Virtual Hub (Section 5).

### 3. The Virtual Hub broker layer

The solution presented in this paper to the problem of the variety and heterogeneity of standards and protocols in the world of geo-information is the introduction of a new mediation component, called Virtual Hub, that extends the SOA archetype. A Virtual Hub in ENERIGIC OD is conceived as a single point-of-access –i.e. a *hub*– to open geospatial datasets (Mazzetti et al., 2015.), which implements some services (i.e. discover, access, distribution, added value, etc.) to facilitate the interconnection between client and server components in a Client-Server architecture and providing all the required mediation and harmonization functionalities among data providers and data users (either developers or end users) (Bigagli et al., 2015.). Virtual Hubs are also *virtual* since they do not physically collect datasets, which are actually kept on their original repository according to a System of Systems approach (Mazzetti et al., 2015).

On the one hand, a Virtual Hub is configured to access a set of data sources that can include individual data sets accessible on the internet, and services compliant with standards like the ones standardized by OGC (catalogue, map, coverage, feature, etc.) or others *accessors* in Fig. 2. On the other hand, clients (end user application and other software) can interact with the VH in order to discover datasets or to access them using the interfaces preferred by their developers (from the powerful interfaces by OGC and ISO to lightweight specifications like OpenSearch *profilers* in Fig. 2). No matter the particular interface a client uses to perform a search or data access query to the VH, the query will be translated and distributed to the appropriate data sources to obtain the desired information and, then, translated again to offer the answer to the application developer in the requested format and via the desired interface. The complexity of developing applications combining different datasets is reduced since the interoperability task is in charge of the broker. ENERIGIC OD VHs also count with web geoportals that handle the interaction with end users using a graphical user interface.

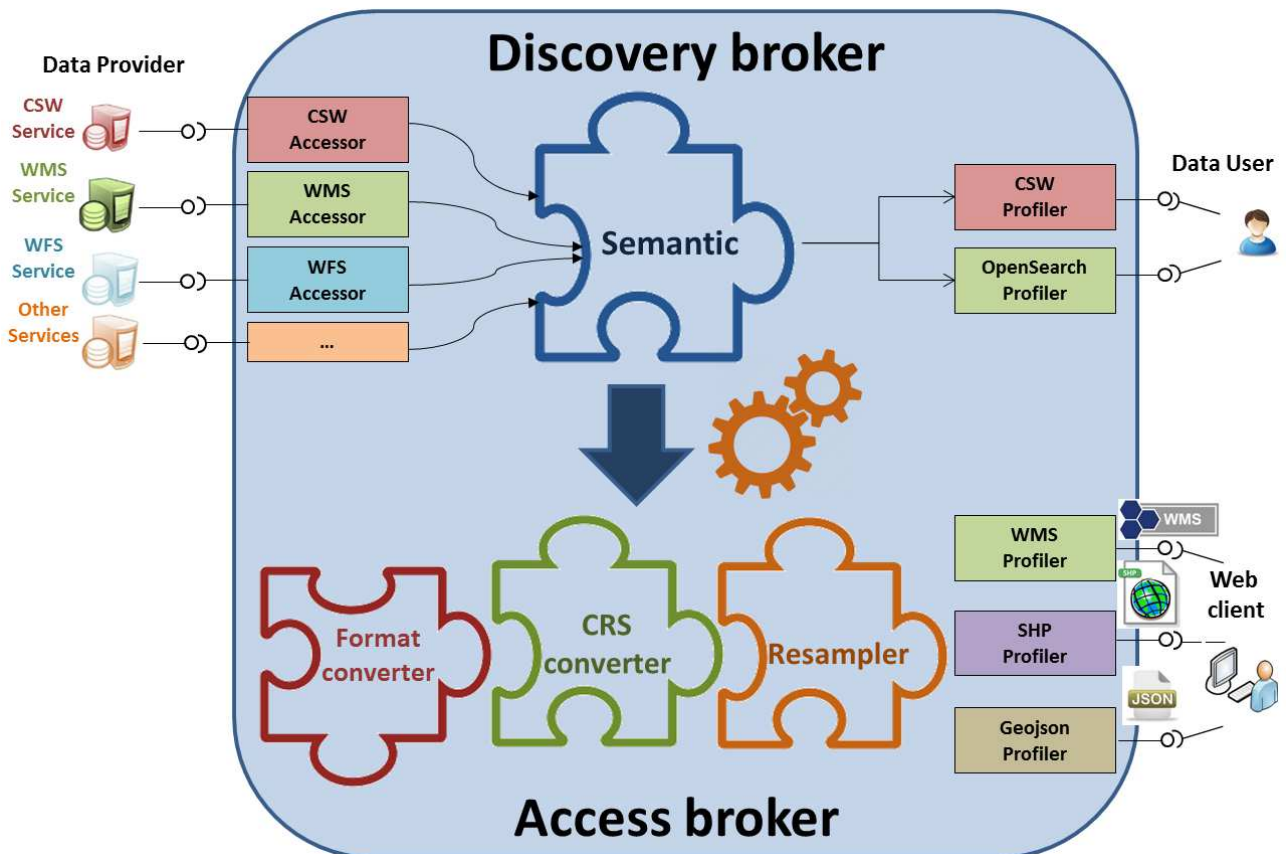


Fig. 2: The broker architecture for heterogeneous data discovery and access.

In particular, the VHs developed in ENERGIC OD project present the following functions and capabilities:

- Distribution of the requests made by clients. As the datasets are kept in their original systems, client requests are distributed across many data providers. One of the major concerns of this approach is, however, waiting time to obtain a query response and the user-perceived performance reduction (Fielding and Taylor, 2002). One of the strategies implemented by the VHs to increase performance is applying a shared caching at the broker to replicate the result of an individual request such that it may be reused by later requests.
- Functionalities to match many server and client protocols, and vice-versa. The supported backend for data sources access includes: OGC standard services (e.g., WMS, WFS, WCS, SOS, CSW), ESRI ArcGIS Geoportal catalogue service, THREDDS, etc. In the frontend, the broker can be accessed by using either the same standards, or other ones like a JavaScript API or a REST API. These APIs expose the brokering functionalities (semantic discovery and dataset access), hiding its complexity.
- Added-value functionalities. The ENERGIC OD VHs implements advanced semantic discovery. In particular, a semantic broker automatically looks for synonyms during a search so that the retrieved results are extended. Two different strategies are used: (i) automatic query expansion: keywords are expanded automatically by the broker by interrogating a set of aligned semantic instruments (typically, controlled vocabularies, thesauri, gazetteers, and ontologies) and (ii) user-assisted query expansion: the semantically-related terms, retrieved from the aligned semantic instruments, are presented to users who are allowed to browse a graph developed according to these terms and select the most pertinent terms.
- Transformation module. This feature is capable of solving for spatial inconsistencies and access data having different coordinate reference system (CRS) and perform data integration in the same CRS.

A great advantage of the brokered approach followed by the VHs is that no common model is defined, and each system can adopt or maintain its preferred interfaces, metadata and data models. The brokered approach allows communities-of-practice to still use their own, preferred specifications, while freeing them of all the mediating between different specifications. Obviously mediation will happen at the lowest common level between specifications but it is generally sufficient for most interdisciplinary applications. The only interoperability agreement is the availability of documentation describing the published interfaces, metadata and data models. No (major) re-engineering of existing systems is required for connecting different data services or data sources. The complexity of interoperability is still there, but it is simply shifted from data/product users and providers towards the brokers. Data users and providers do not have to make their clients and server compliant with specifications anymore but the broker is in charge of handling all the complexity.

Additionally, ENERGIC OD Virtual Hubs can partially solve the *variety* issue of Big Data, since heterogeneous geospatial data sources are accessed in a seamlessly way. In particular, the heterogeneity in data formats, standards and protocols is solved by the VH allowing a higher interoperability between services and datasets. VHs are also able to face the Big Data issue of data *volume* allowing to access potentially unlimited datasets.

#### 4. Related Work



1 The Virtual Hub broker described in section 4 and used in the development of the applications presented in  
2 section 5 was developed by the ESSI-Lab of the Italian Consiglio Nazionale delle Ricerche as part of the GI-  
3 suite. The GI-suite is a brokering framework targeting interoperability of heterogeneous systems in the  
4 Geoscience domain (Boldrini et al., 2015). It implements discovery, access and semantic brokers (Nativi et  
5 al., 2013). The brokering takes place between a set of heterogeneous publishing services (more than 40  
6 publishing standards and implementations are supported) and a set of heterogeneous consumer  
7 applications (more than 15 standards and implementations are supported on this side).  
8

9  
10 It has query expansion capabilities implemented by accessing to semantic assets (vocabularies, thesauri,  
11 ontologies) stored in a knowledge base. It is capable to integrate quality information provided by data  
12 producers, and feedback from users, together with access brokering (including transformation services to  
13 interconnect access clients and servers).  
14

15  
16 It has already been used in several projects and has been improved through them (EuroGEOSS, GEOWOW,  
17 ENVIROFI, GeoViQua, MEDINA, ODIP and ENERIGIC OD) and it is characterised by aiming at facilitating  
18 multidisciplinary interoperability (Nativi et al., 2012; Giuliani et al., 2017). It is the basis of the GEO  
19 Discovery and Access Broker (DAB) component of the GEOSS Common Infrastructure (GCI) GEOSS (Nativi et  
20 al., 2015).  
21

22  
23 Other brokering approaches in the context of GI information can also be found in the literature. As the  
24 harmonization and standardization process are not a universally applicable solution, the System of Systems  
25 approach is prevalent among this works as the selected means for facilitating data sharing and reuse  
26 among different communities. Under this approach, brokering or mediation allows for system-to-system  
27 communication among the systems already in place with no need of making any modifications to them  
28 (Béjar et al., 2009; Nativi et al., 2014). However, and regarding its application to the development of the  
29 applications on historical cartography presented in section 5, the other identified frameworks (that are  
30 discussed below) are designed for the brokering of specific types of scientific data or rely for their  
31 mediation capabilities on explicit data models. This fact would make its use in the context of historical  
32 cartography impractical, despite the fact that many of them are powerful tools, very well-suited for the  
33 domains they target.  
34

35  
36 For instance, the gCube system<sup>1</sup> is an open-source software toolkit designed to enable the creation and  
37 operation of Hybrid Data Infrastructures: infrastructures built as a “system of systems” that integrates  
38 other infrastructures including grid and cloud, services and information systems (Assante et al., 2014).  
39 gCube can serve the needs of diverse communities of practice by providing each with web-based working  
40 environments where groups of scientists can transparently and seamlessly access shared sets of resources  
41 (Assante et al., 2016). gCube can be defined as a broker, as it is designed to abstract over a variety of  
42 technologies such as data, process and resource management. gCube users can seamlessly access to  
43 species data, geospatial data, statistical data and semi-structured data from diverse data providers and  
44 information systems (Candela and Pagano, 2015). As with the GI-Suite, these services can be exploited both  
45 via web-based graphical user interfaces and web-based protocols for programmatic access, e.g., OAI-PMH,  
46 CSW, SDMX.  
47

48  
49 gCube particular focus is in scientific open data, with a particular emphasis on species and biodiversity data,  
50 for which it has specific tools and components (Candela et al., 2015). Regarding geospatial data, gCube is  
51 equipped with services capable of generate Spatial Data Infrastructures compliant with OGC standards. In  
52

53  
54  
55  
56  
57  
58  
59  
60  
61 <sup>1</sup> <http://www.gcube-system.org/>  
62  
63  
64  
65

1 particular, it offers a catalogue service enabling the seamless discovery of and access to every geospatial  
2 resource registered or produced via gCube services, but again, focused on biodiversity scientific data.  
3 gCube provides more advanced functionalities and acts more like a broker than most science gateways  
4 (Candela et al., 2015; Shahand et al., 2015), but it is still mainly focused in scientific data and adapted to  
5 work with it.  
6

7 The Environmental Research Division's Data Access Program (ERDDAP) data broker<sup>2</sup> provides scientists with  
8 a simple, consistent way to download subsets of gridded and tabular scientific datasets in common file  
9 formats and make graphs and maps. ERDDAP is an Open Source Cloud based technology developed by the  
10 Environmental Research Division of NOAA with the specific purpose of making it easy for software  
11 developers to access scientific data and, in particular, gridded data such as that outputted by  
12 oceanographic and atmospheric numerical models as NetCDF (Network Common Data Form) files. ERDDAP  
13 works as a broker or mediator by converting data requests to any resource listed in its catalogue into the  
14 language used by remote data server and by converting original data formats into others more familiar to  
15 the developers.  
16  
17  
18  
19

20 Unlike the GI-Suite, ERDDAP relies in offering only the OPeNDAP standard to request data and it is mainly  
21 focused, as gCube, in the scientific domain. Indeed, all ERDDAP deployments found by the authors deal  
22 with marine and oceanographic data.  
23  
24

25 The CIDS product suite<sup>3</sup> consists of a set of services, applications, software components, application  
26 programming interfaces (APIs) and management and development tools for the management, integration,  
27 and development of heterogeneous information systems with a special focus on interactive geospatial  
28 systems. It provides a distributed integration platform, which is particularly useful for workflows that need  
29 a combination of information and processes from different source systems such as GIS systems, relational  
30 databases, and simulation models. In this way, it already provides and supports a considerable number of  
31 functionalities required for complex geospatial information systems, including user management and  
32 access control, search and discovery of relevant information and advanced interactive 2D visualisation  
33 (OGC WMS and WFS clients).  
34  
35  
36  
37

38 As the GI-Suite and gCube, CIDS is particularly suited for solutions which have to be built across existing  
39 heterogeneous information systems, which may be under control of different organizations (Dihé et al.,  
40 2015). Unlike gCube and ERDDAP, it is not primarily dedicated to scientific data, although its main focus is  
41 not brokering data, but managing process workflows. As a broker, it has its stronger capabilities as a  
42 discovery broker. Another difference with the GI-Suite, that is domain-model agnostic, CIDS data access  
43 brokering is based on the creation of specific domain models, and tools for conversion and schema  
44 mapping facilities may be necessary (Denzer, 2014). CIDS is being applied mostly in the water science  
45 domain with the help of specific domain models (Torres-Bejarano et al., 2011; Hell et al., 2013).  
46  
47  
48  
49

50 Besides the aforementioned frameworks, that can truly be considered brokers or mediators under the  
51 system-of-systems approach, there exist in the literature many other initiatives and technologies that try to  
52 facilitate either the discoverability of data or the accessibility to them through virtual nodes (i.e., through  
53 servers that do not store themselves the data). The Catalogue Services for the Web Mediated (CSWM),  
54 developed in the NETMAR project (Lassoued, 2012; Leadbetter et al., 2014), is an example of the former.  
55 The CSWM is an OGC CSW semantically enabled does not harvest or index the metadata records of the  
56  
57  
58

---

59  
60 <sup>2</sup> <http://coastwatch.pfeg.noaa.gov/erddap/index.html>

61 <sup>3</sup> <https://www.cismet.de/en/cidsReadme.html>  
62  
63  
64  
65



catalogue nodes. Instead, it rewrites queries into the languages and interfaces supported by the catalogue nodes and executes them on the fly, collecting the answers from the different nodes and sending them back to the user. It is focused on enhancing discovery in a distributed environment (so it is a discovery broker), but not into easing access to the discovered data. Additionally, as with other frameworks, its prime target is marine environmental data. As an example of the latter (virtual nodes that provide accessibility), we can cite the work by Regueiro et al. (2015), that applies the mediator/wrapper architecture for the virtual integration of heterogeneous observation data sources through a Sensor Observation Service (SOS) standard interface.

## 5. Historical Map web Applications

This section presents two applications dealing with historical maps built on top of a Brokering Virtual Hub presented in section 3. In particular, (i) GeoPan and (ii) Zaragoza Histórica applications are presented, focusing on advantages of using the presented VH solution. Even if the two applications deal both with historical maps, they present significant differences both at technical (e.g., libraries, development environment, data formats, etc.) and non-technical (e.g., user addressed, user requirements, etc.) aspects. While GeoPan, addressed to professional users, highlights the evolution of the territory by integrating information coming from historical maps, current cartography and satellite data, with a particular focus on tracking the riverbed changes occurred across time, Zaragoza Histórica is addressed to the general public by providing a really easy-to-use web tool to visualize and compare historical cartography of a city. On the technical side, the two applications are using two sets of different specific functionalities of the VH: JavaScript API for GeoPan and the REST API for Zaragoza Histórica. This clearly exemplify the flexibility of the developed approach to address different requirements.

Next section jointly presents the data used by both applications, the digitization and georeferencing processes applied to the involved historical cartography and discussed the accuracy obtained. Section 5.2 describes with more detail the GeoPan application and some applications are discussed in subsection 5.2.1. Section 5.3 presents the application Zaragoza Histórica while subsection 5.3.1 focuses more on technical implementation aspects.

### 5.1 Data involved, georeferencing and accuracy

Both applications deal mainly with historical cartography, but a variety of data sources, both current and historical datasets have been used in order to achieve the desired functionality of each application. A list of data is presented in Table 1 and Table 2.

Map	Typology	Century	Scale	Covered area	Application
Theresian Cadastre	Large Scale	Second half XVIII	1:2000	15 Municipality in Lombardy region (among them: Milan, Monza, Treviso, Gorgonzola, Veduggio al Lambro)	GeoPan

Lombardo-Veneto Cadastre	Large Scale	Beginning XIX	1:2000 attachments 1:1000	50 Municipality in Northern Italy (among them: Milan, Monza, Tremezzo, Gorgonzola, Vedano al Lambro, Verona, Padova)	GeoPan
New Lands Cadastre (Cessato Cadastre)	Large Scale	Late XIX – beginning XX	1:2000 attachments 1:1000	50 Municipality in Northern Italy (among them: Milan, Monza, Tremezzo, Gorgonzola, Vedano al Lambro)	GeoPan
ITN military maps	Medium Scale	Second half XIX	1:50.000 1:25.000	Northern Italy (Lombardy and Veneto)	GeoPan
Carte di Monza e Brianza (Brenna)	Medium Scale	Second half XIX	1:25.000	Lombardy (Milan, Monza, Como, Lodi)	GeoPan
Maps of Italian cities (Vallardi)	Large Scale	Second half XIX	1:5000	Italian Cities (among them, Firenze, Modena, Ferrara, Reggio Emilan)	GeoPan
Zaragoza Municipal Archive maps	Large-Medium Scale	XVIII, XIX, XX and XXI (19 maps from 1712 to 2004)	1:10000 1:1250	Zaragoza City	Zaragoza Histórica

Table 1. List of historical data used in the presented applications that needed georeferencing.

Data	Typology	Source	Equivalent scale	Reference system	Application
Rivers and streams	Vector data (line)	Local SDI (Geoportale Lombardia)	1:10000	WGS84/UTM32	GeoPan
Wells	Vector data (point)	Local SDI (Geoportale Lombardia)	1:10000	WGS84/UTM32	GeoPan

Seismicity scenario (Microzonazione sismica)	Vector data (polygon)	Local SDI (Geoportale Lombardia)	1:10000	WGS84/UTM32	GeoPan
Orthophoto 1975	WMS	Local SDI (Geoportale Lombardia)	1:15000	Roma40/Ovest (Monte_Mario_Italy_1)	GeoPan
Orthophoto 1998	WMS	Local SDI (Geoportale Lombardia)	1:15000	Roma40/Ovest (Monte_Mario_Italy_1)	GeoPan
Sentinel data (S2A) used to extract thematic layers (e.g., soil moisture, NDVI, etc.)	Raster data	Esa Portal	-	WGS84	GeoPan
Georeferenced historical urban maps	Raster data / OGC:WMS (via the Virtual Hub)	Zaragoza City Council / VH- ES	1:5000	WGS84/UTM30	Zaragoza Histórica
Current base cartography of Zaragoza City	OGC:WMTS	Zaragoza City Council (Zaragoza SDI)	1:5000	WGS84/UTM30	Zaragoza Histórica

Table 2. List of Open Data used in the presented applications.

Those datasets are published by different data providers with various data format and interface (Table 3).

Data	Provider	Data format/service
Georeferenced historical territorial and cadastral	Gicarus Lab	OGC: WMTS
Georeferenced historical urban maps	Zaragoza City Council	OGC:WMS (via the Virtual Hub)
Historical orthophotos	Local government – Lombardy	OGC:WMS
Geological Open Data layers	Local government – Lombardy	ESRI shapefile
Satellite data downstream services	Gicarus Lab	OGC:WCS
Current base cartography	Zaragoza City Council	OGC:WMTS

Table 3. Data used in the GeoPan APP with their own providers and data format/service used by the publisher.

These maps, stored in different archives, had to be digitalized in order to fully exploit them in the applications. The responsible archives had already performed the digitalization in their premises and with their own resources, producing sets of high resolution files (for instance, in TIFF format in the case of the Archive of Zaragoza). However, these digital maps cannot be used directly in the application, as they must be georeferenced first. Transformation to a target coordinate system is the main goal of georeferencing the historical maps. Two methods of georeferencing may be distinguished: transformation between two coordinate systems where transformation parameters are known, and transformation using identical

1 ground points where transformation parameters are unknown. The techniques for georeferencing depend  
2 on the characteristics of the historical maps: the quality of the geodetic grids, the map projections,  
3 generalization methods, etc.

4 If the original projection is known and precise, the map sheets can be projected using unique  
5 transformation parameters, with projection parameters including a central meridian and the shape of the  
6 map sheet (four points at the edges). Projection error rises if a meridian convergence is not taken into  
7 account. This method can be applied on a small scale for low accuracy (or coarse) transformation. However,  
8 in many cases there are many uncertainties in the original maps, with respect to: map projection, central  
9 meridian, scale distortion, survey technique used for the realization of the map, etc. In this case the  
10 georeferencing can be carried out without known parameters by: (i) using reference data sets, (ii) finding  
11 identical points, and (iii) applying an appropriate transformation method. This is the choice adopted in this  
12 work. In particular, the georeferencing of historical maps was carried out in a two-step way:

- 13 ● Georeferencing of historical maps to a contemporary coordinate system;
- 14 ● Mosaicking of the individual map sheets into a unique one.

15 In contrast to other approaches in literature (Podobnikar, 2009; Brovelli and Minghini, 2012) that are  
16 operating in the opposite way (i.e., first mosaicking and then georeferencing), we operated this choice  
17 since the different maps sheets may have different deformations each other and/or different accuracies.  
18 For example, highly vegetated or mountainous areas may have lower accuracies with respect to urbanized  
19 areas and by putting together those two sets of maps the overall accuracy of the final one can be  
20 worsened. On the other hand, by processing and georeferencing each image separately first and then using  
21 boundary points to stitch them together the final result is a global estimation that avoids error  
22 accumulation, whereas image-to-image correspondences plays as tie points and strengthen block geometry  
23 (Barazzetti et al., 2014). More specifically, after setting Ground Control Points (see below) on both an old  
24 and a new map, a global transformation (generally a polynomial or rubber sheet, see also Brovelli and  
25 Minghini, 2012) is used to transform points on the old map to the new map and vice versa. Parameters for  
26 the transformation are calculated with the ‘least-squares-method’, resulting in minimum errors between  
27 point sets. The polynomial order 1 is appropriate for transformation between two projections where great  
28 distortion is not present (useful for transformation between newer maps). The polynomial order 2 is  
29 appropriate for transformation between geographical coordinates (or sphere) and projected data, or for  
30 locally systematically distorted data sets in larger areas. The rubber sheeting method could be more useful  
31 on more highly distorted areas, but it requires a high number of identical points and unpredictable  
32 distortions could appear around areas of deficiency. For this paper we have used mainly the affine-  
33 transformation (six parameter-transformation) which only uses scaling, rotation, translation in x and y, and  
34 shear in x and y and therefore leaves the historical map as original as possible during overlay with a modern  
35 map. Thin plate spline was used in the case of maps presenting large deformations with respect to the  
36 current cartography. This was mainly used for Zaragoza maps, where the older maps presented such  
37 deformations. As mentioned before, in order to transform the historical map, identical points on both the  
38 old and new map (Ground Control Points, GCPs) need to be identified, a somewhat delicate task (Guerra,  
39 2000; Balletti, 2006). Potentially the best points are the measured ones: churches or towers. To locate  
40 appropriate points, it is important to note the “style” of a particular cartographer and cartographic  
41 elements, including the level of generalization of the maps. Additionally we should be very attentive to  
42 possible landscape changes (e.g., riverbed changes).

43 Therefore, attempts were made to:

- distribute the GCPs as uniformly as possible over the entire map
- use various features as GCPs
- use unambiguous GCPs as road crossings, fortresses, building corners, etc.
- not use features that change over time (e.g. landmarks)
- set a large number of GCPs.

The chosen identical points in the used for georeferencing were: churches (especially those on the peaks), characteristic bridges, road crossings, towers, and occasionally, distinguishable crossroads (for example, Fig. 3a shows the fixed set of 21 GCPs used for georeferencing each one of the maps in the Zaragoza Histórica collection) and Fig. 3b shows 18 GCP used for georeferencing a map sheet of the Lombardo-Veneto cadaster in the central part of Milan (it can be easily recognized the cathedral) and the corresponding points in the OSM map. However, in some small-scale maps, towns were only depicted using church or point symbols. In that case, the ‘centers’ of present-day towns were defined by the main churches which were used as GCPs.

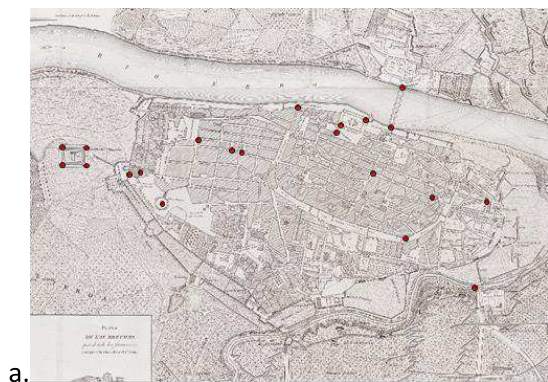


Fig. 3: GCP selection: (a) selected, fixed GCPs for Zaragoza Histórica maps; and (b) selected GCPs for a map sheet of the Lombardo-Veneto cadaster in the central part of Milano (left) and corresponding points in current OSM cartography (right).

Quality assessment was performed in the georeferenced images with the aim of the is to analyse the quality and errors that can occur in the MS1 maps by statistical tests (and visualisation) of the possible errors. The following positional accuracy parameters were evaluated:

- Root mean square error (RMSE) which represent an overall quantification of the planimetric accuracy. In particular RMSE is defined as:



$$RMSE = \sqrt{\frac{\sum v^2}{N}}$$

Where  $v$  is the residual (i.e., the vector distance in meters between the actual points in the old map and the place they would be according to least-squares estimation, see also Fig. 4) and  $N$  is the number of selected control points. The result should be interpreted as the average distance between a randomly chosen point on the map and its actual position in case the old map is actually as accurate as the modern map. Statistics on GCPs residuals are a measure of the transformation precision. Since thin plate spline is not producing residuals this measure can be performed only for affine transformation of maps.

- Averaged error (AE) was calculated as differences between manually selected control points on georeferenced maps in comparison with reference data sets; it can be considered as an indication of transformation accuracy.

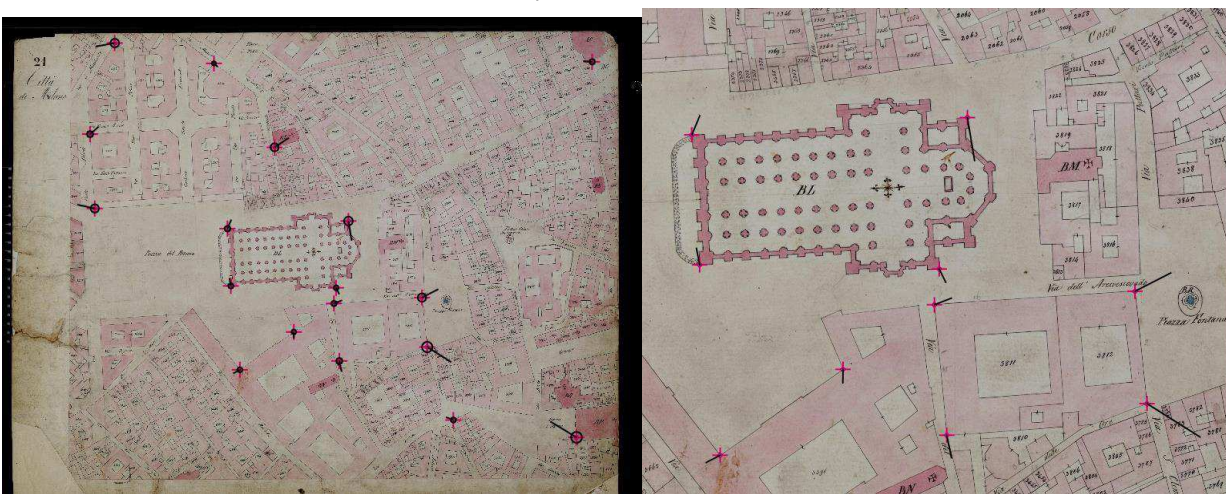


Fig. 4: Residuals after georeferencing a map sheet of the Lombardo-Veneto cadaster in the central part of Milano: the entire map sheet (left) and a detail (right). Vectors representing residuals are magnified of a factor 10.

Applying the classical scale values (<5000 as large-scale, 5000 to 50,000 as medium-scale, and >50,000 as small-scale). The majority of the georeferenced maps falls into the large-scale and the medium-scale category (Table 4).

	Large-scale	Medium-scale	Small-scale
<b>GeoPan</b>	68	52	10
<b>Zaragoza Histórica</b>	19	-	-

Table 4. Number of maps georeferenced subdivided according to scale values principle.

Since the two presented applications (GeoPan and Zaragoza Histórica) are focusing on different geographical areas (Italy and Spain respectively) and historical data were georeferenced by using different models (affine for GeoPan and thin-plate spline for Zaragoza Histórica), quality evaluation is performed independently.



The first tests that conducted concerned the RMSE (only for GeoPan) of the series of maps. In this case the analysis focused only on large-scale and medium-scale maps since small-scale maps represents a too small sample for any type of reliable result. The boxplots in Fig. 5a show a large difference in RMSE between the maps of different scales, as expected, (i.e., median values differ significantly). Similar consideration can be drawn also for the AE. A comparison between RMSE and AE shows both in the median values and quintiles a similar behavior highlighting both the precision of the transformation and the accuracy of the output map. Also the 'distortion maps' (Jenny, 2010) generally present a regular grid (Fig. 6) showing the suitability of the affine transformation model in the majority of cases, especially with large scale maps. Considering Zaragoza Histórica only AE was evaluated (Fig. 5b) since, as previously discussed thin-plate spline, does not provide residuals. In this case the evaluation was performed on a selected set of historical maps. The selected maps are representative of different epochs, ranging from 1712 to 2004. In particular, we can observe that maps belonging to the XVII and early XIX century presents an average error significantly larger with respect to the average one. This can be due to the fact that historical maps presents an highly irregular deformation. Thin-plate spline can be effectively used to smooth this effects (Affek 2013). However, some residual deformation may persist.

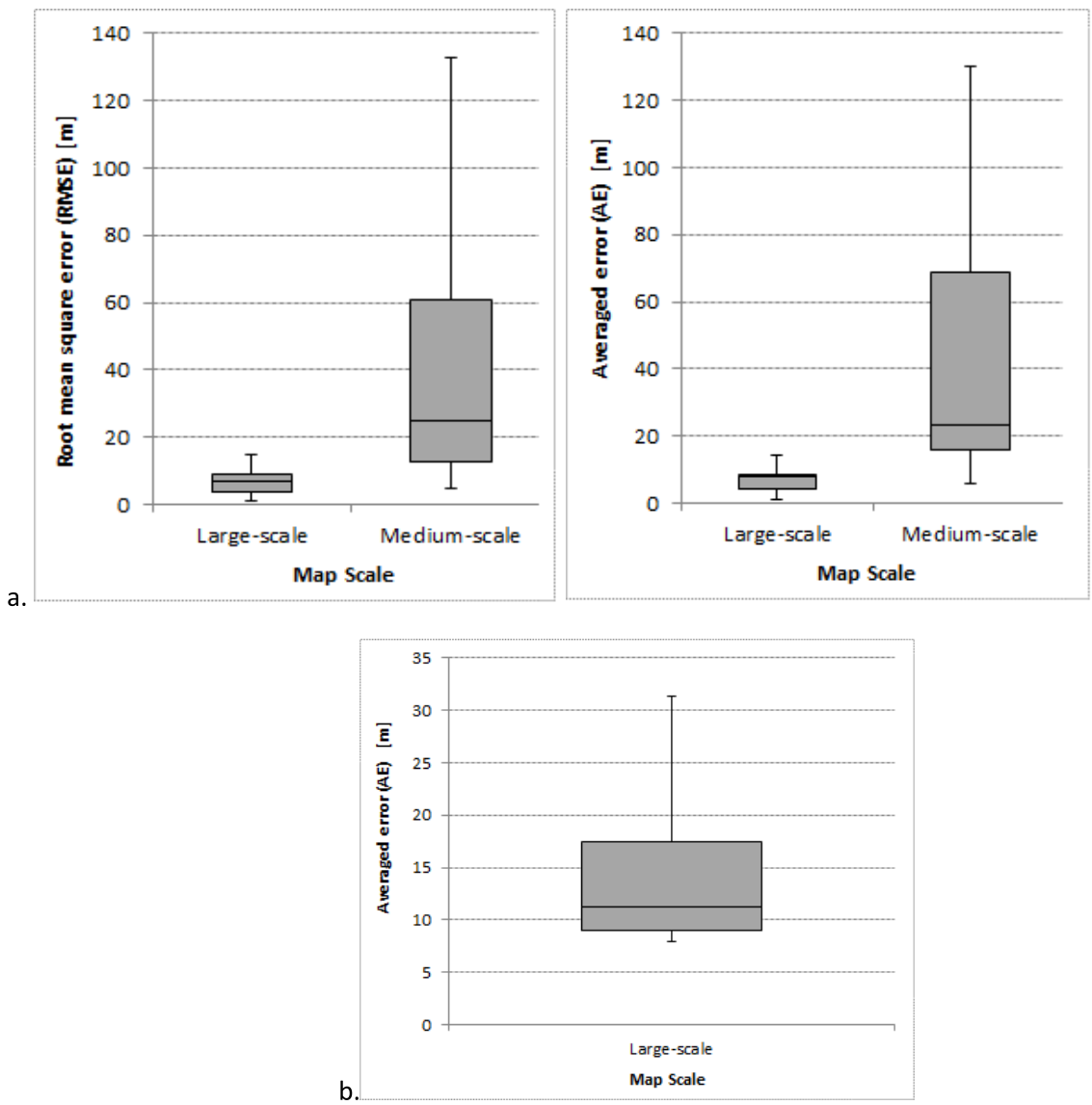


Fig. 5: Historical maps boxplots: (a) boxplot of the RMSE (left) and AE (right) of large- and medium-scale maps for the GeoPan application; and (b) boxplot of AE for the Zaragoza Histórica.

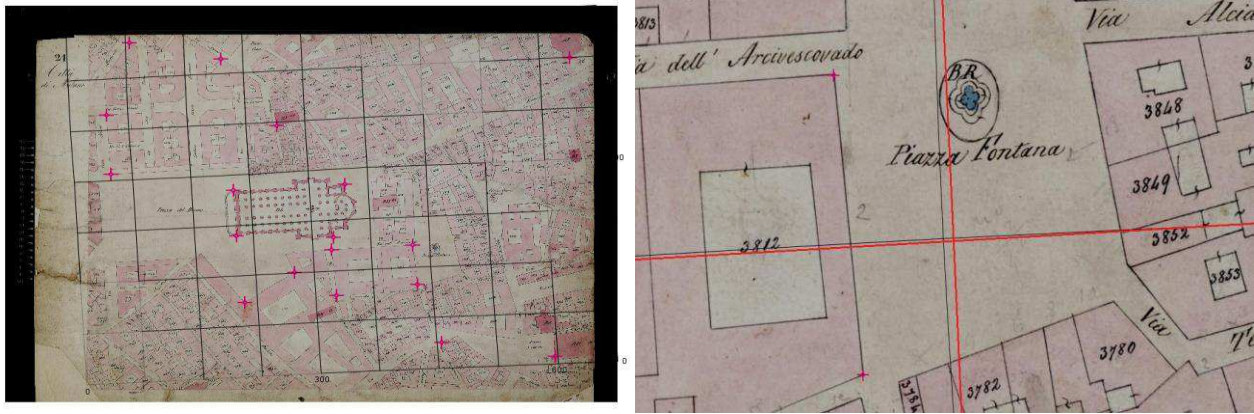


Fig. 6: Distortion map for a map sheet of the Lombardo-Veneto Cadaster in the central part of Milano: the regularity of the grid shows the suitability of the affine model: the overall map sheet (left), and a detail showing the area with larger distortion (right). In black distorted lines and in red undistorted ones.

Next, the evolution of the RMSE (for GeoPan) over time was analyzed (Fig. 7). Overall map accuracy increased over time (decreasing RMSE). However, it is best to look at the different map scales separately. Results point to increasing map accuracy over time for medium- and small-scale maps, for small-scale maps the  $R^2$  value (0.875) quite high interrelation of time and accuracy. This can be particularly appreciated focusing on the three cadastral map series. In particular, RMSE values allows appreciating an increase in precision of the Lombardo-Veneto cadastral maps (mid 19th century) and their updates of 1898 with respect to the earlier Theresian Cadastre (mid 18th century). This result can be explained taking into consideration the improvements of map surveying techniques from the 18th to the 19th century. For example in Theresian cadaster maps no geodetic framework was considered and the problem of projecting the Earth's surface on the plane was ignored. The updates of 1898, whose map sheets had been drawn on the basis of their corresponding Lombardo-Veneto ones, showed slightly better precision and accuracy compared with the Lombardo-Veneto ones. This may be due to the lower deformation of paper supports over time of this last series of maps.

For medium scale maps only 58.5 % of the variation can be explained by the independent variable time (see also Grabowski and Gurnell, 2016). The remaining portion of the variation is associated to other factors. In particular, some medium-scale maps have exceptionally larger RMSEs than expected providing a scattered behavior to the data. More specifically military maps tend to be more accurate than other ones (e.g., Brenna maps) since all of them are in the first quintile. Due to the limited number of maps analyzed in the case of Zaragoza Histórica it was not possible to derive meaningful results addressing the link between year of publication and map fitting with current cartography.

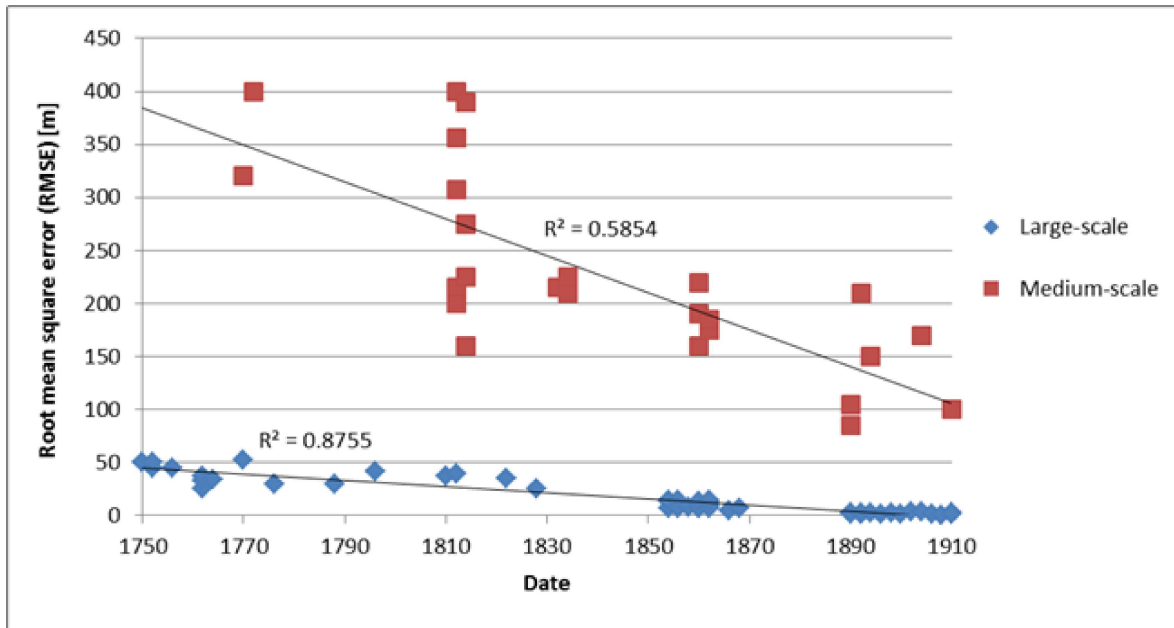


Fig. 7: RMSE over time for large- and medium-scale maps for the GeoPan application.

## 5.2 GeoPan

GeoPan deals with multi-temporal and multi-source geospatial data query, visualisation and consultation by using a webGIS geo-portal concept as a gateway to information. During the development of the web application, choices were made by paying attention to the needs and peculiarities of the project: (i) visualisation and management of large dimension raster maps, (ii) organization of cartographic resources for easy consultation, and (iii) interface with a visual impact able to attract and to clearly communicate the information requested (Fig. 8).





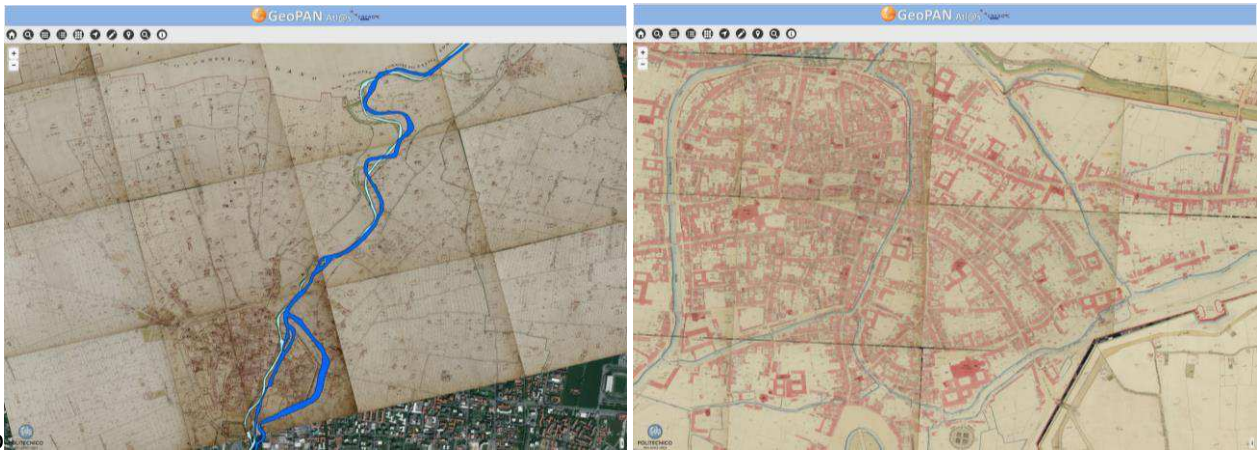
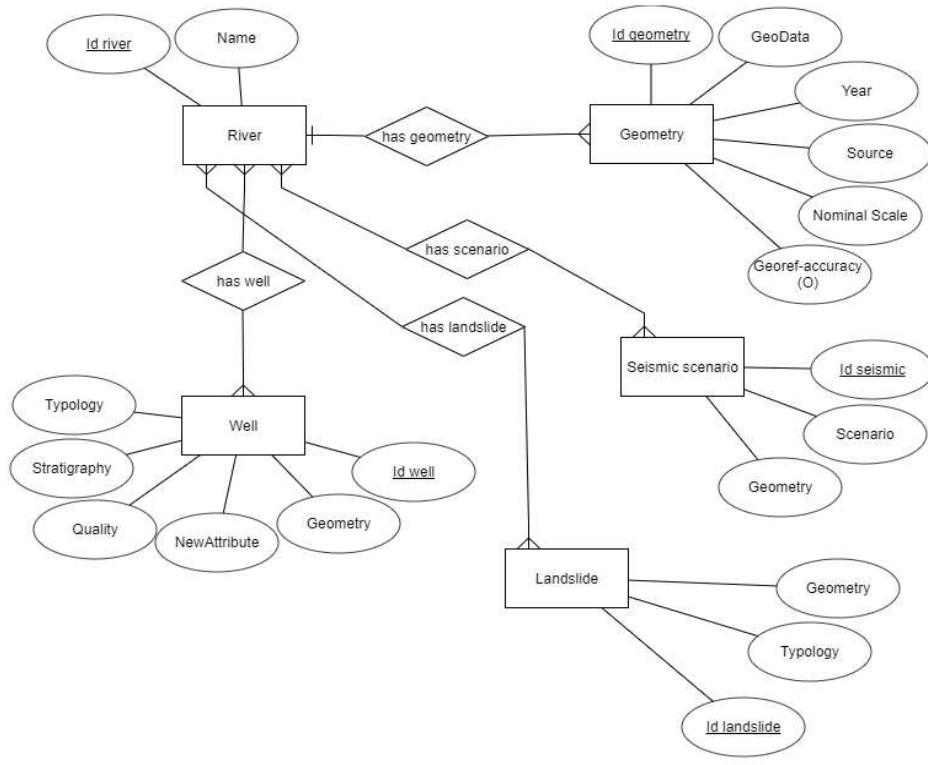


Fig. 8: User interface of GeoPan for mobile (a) and desktop (b) devices.

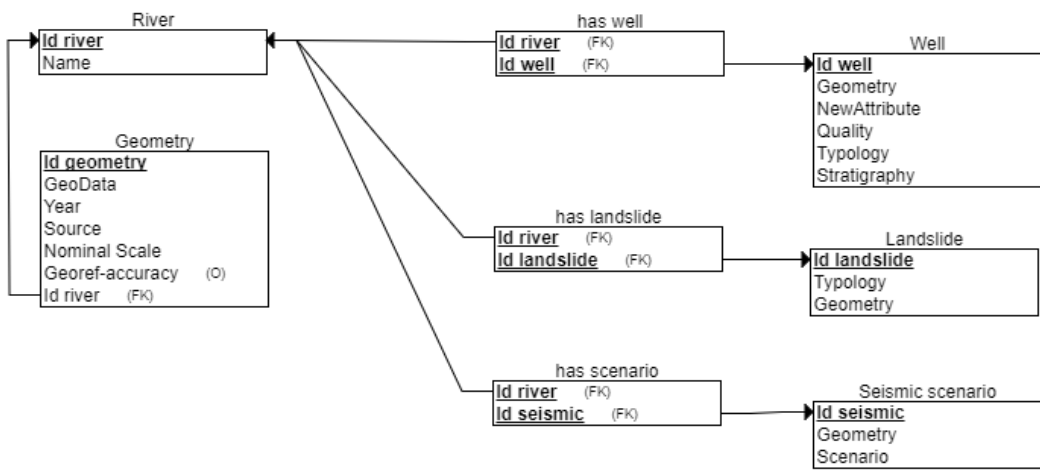
The application is mainly devoted to highlighting the evolution of the territory by integrating information coming from historical maps, current cartography and satellite data. In particular, GeoPan APP is focusing on tracking the riverbed changes occurred across time in Lombardy (North Italy). The presence of former riverbeds represents a vulnerability both in case of seismic and flooding events. Indeed, the soil in correspondence of former riverbeds is generally looser than the surrounding one determining some amplification of the seismic waves. In addition, those areas are more prone to be flooded in the case of intense rainfall events.

The different vector datasets (Table 2) were organized into a spatial database. In particular, taking into consideration the different OD provided by the local SDI, five main entities, each one with geometrical attributes, were identified. The conceptual model of the database is provided in the form of an Entity Relationship Diagram (ERD) in Fig. 9a, where the entities, the relationships between them and the cardinality of the relationships are represented.

The database ERD provides a peculiar complexity due to the temporal component. For example, each specific river may undergo different reshaping of the riverbed during the time and when this information is considered, the number of possible relationships between the entities increases. The majority of relationships defined in the model have N:N (many-to-many) cardinality. The only relationship having 1:N (one-to-many) cardinality is the one connecting "River" and "Geometry". Indeed, one river may have a different geometry of the riverbed (depending on the epoch considered) while one riverbed geometry is associated only to one river. Lastly, the logical model of the database (Fig. 9b) was designed, which defines the attribute data for each entity in the database. For the choice of the RDBMS (Relational Database Management System) to be used for the system, we opted for a platform based on PostgreSQL, with PostGIS extension for spatial data management. The choice was made mainly for its open source nature (Fig. 10b).



a.



b.

Fig. 9: Database of vector elements: (a) ERD of the database showing entities (boxes), attributes (ellipses) and relationship (diamonds) and (b) logical model of the database.

The advantage of developing GeoPan on the top of the Virtual Hub is given by the fact that instead of connecting the application to the different services publishing the requested data a single connection is established to the VH by using a set of simple Web APIs. The mediation and the connection with the different services is in charge of the VH reducing the time and the effort for the development of the application. Different FOSS (Free/Open Source Software) solutions were considered for both server and client-side implementation. In particular, the historical maps of the areas are published, once performing their georeferencing, by using a server-side instance called Gicarus Lab<sup>4</sup>. The server-side software used is

<sup>4</sup> <http://geoserver.atlas.polimi.it/geoserver>

GeoServer<sup>5</sup>, a powerful Open Source platform for publishing spatial data and interactive mapping applications on the web. GeoServer allows the visualisation of produced data through the use of OGC standards like Web Map Service (WMS, WMTS and WCS). The time needed for loading and navigating these data represents the biggest challenge in accessing and displaying large raster maps in an efficient way, the largest map published by the service has an original size of 200GB. For this reason, historical maps are released by using WMTS which is an OGC standard protocol for serving pre-rendered georeferenced map tiles on the Internet. Indeed, for most WMS services it is not uncommon to require 1 or more CPU seconds to produce a response. In the case of massive parallel use, such CPU-intensive service is not practical. To overcome the CPU intensive on-the-fly rendering problem, WMTS allows for using pre-rendered map tiles. The general architecture of the GeoPAN APP is presented in Fig. 10a. An architecture with the VH gives a simpler access to geographic OD, giving a single point of access with a standard protocol and interface. In particular, the Virtual Hub Layer is used to collect and access different services and cartographic information provided by local SDI, historic maps provider and datasets (historical and thematic layers) published by the Gicarus Lab Server (Fig. 10b).

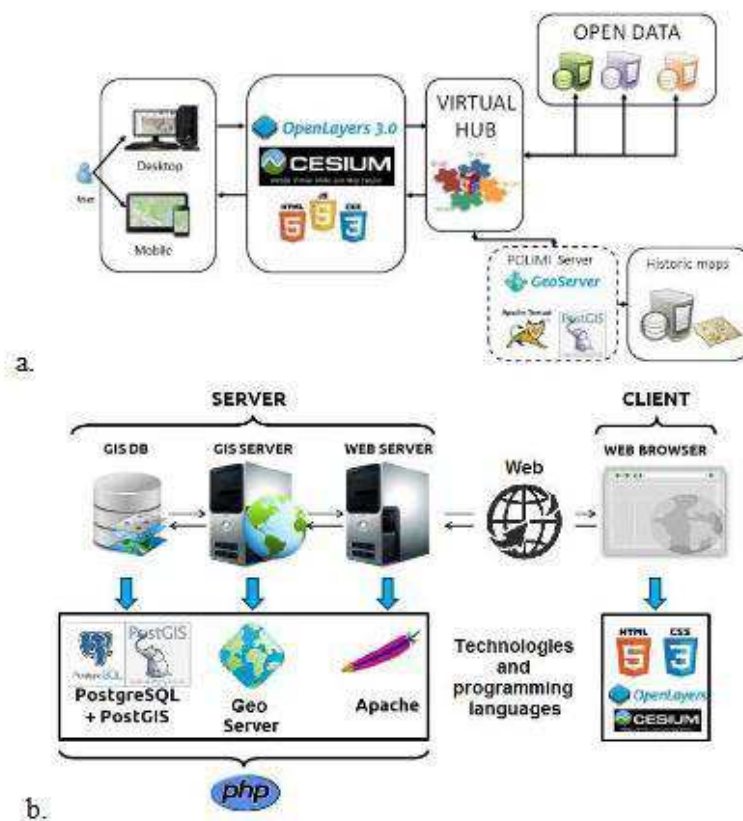


Fig. 10: GeoPan deployment diagram (a) and server-client architecture (b).

To develop the client-side of the system of the GeoPan application, different FOSS libraries were used. In particular, front-end/interface is developed in a classic HTML/CSS/JS approach. It uses WebGL for hardware-accelerated graphics, and is cross-platform, cross-browser, and tuned for dynamic-data visualisation, it also is supported by Android, iOS and Windows tablet. In addition, the application is released also as an .apk file, developed by using the Apache Cordova technology (PhoneGap), for Android devices. The GeoPan application was developed using a set of Open Libraries (delivered under the Apache 2.0 license, which means they are free for commercial and non-commercial uses):

<sup>5</sup> <http://geoserver.org>



- PhoneGap (Apache Cordova): PhoneGap is a free and open source framework that allows to create mobile apps using standardized web APIs for a very wide range of mobile platforms.
- jQuery: the jQuery 1.10.2 library is used to develop the front-end (UI and application behaviour). Some jQuery plugins are added such as blockUI, numeric, and zebra datepicker in order to enhance the user experience.
- OL3-Cesium: OL3-Cesium is a JavaScript library for creating 3D globes and 2D maps in a web browser without a plugin. This library is used as a support to manage layers returned from queries performed by users.
- giapi-1.2.9-beta<sup>6</sup>: Virtual Hub JavaScript library. Required to interface the application with the Virtual Hub.

The GeoPan web client consists of two parts: a layer menu and a main map panel. The layer menu itself is additionally divided into two parts: the first one, that is devoted to user with basic IT skills, allows users to view and select all the georeferenced historical maps which are ordered into separate trees, while the second one, devoted to users with programming skills, allows direct interaction with the Virtual Hub formulating queries and displaying results.

### 5.2.1 GeoPan in action

As previously described, GeoPan is aimed at combining historical maps with local current cartography to track land and territorial changes during the centuries. Specifically, the application is focusing on riverbed since those areas are more vulnerable to both hydrological and seismic risk. The aim of this paragraph is to present two examples of riverbed change tracking by using GeoPan in the Valtellina area (Northern Italy) and to show its potential expansion in other areas and in other countries other than Italy. The first case study is the area of the so called "Pian di Spagna". The Pian di Spagna origins are connected with the transportation of sediments by the Adda river around XVI century. In that period the amount of sediment transported by Adda caused the outcrop of a broad plain. The newborn Pian di Spagna became thus a vast area of marshes and swamps, wild and malaise, tormented by malaria. In 1829 the Austrian government decided to channel and rectify the final part of the Adda, with the aim of reclaiming the land. In particular, in 1829 Giuseppe Cusi had the task of "drafting a new rectilinear riverbed for the Adda and a canal for the emissary of Lake Mezzola that would allow navigation." So, in the surroundings of Piantedo, the riverbed was diverted to a channel 4.2 km long and 140 m wide (Fig. 11a-top). A large swamp area was dried and cultivated. Following these works in 1858, the Adda abandoned forever its old riverbed and moved from the emissary of Lake Mezzola to Lake Como. However, the presence of the former riverbed makes this area more prone to seismic risk amplification. Fig. 11b-top shows the link between areas characterized by amplification of the seismic effects and the former Adda riverbed.

The second case study is the area of the so called "Pian della Selvetta". The plan of the Selvetta is a vast territory occupying the middle part of Valtellina between St. Pietro Berbenno and the Masino stream in the municipality of Ardenno. This territory has always been a swampy territory due to several meandering, in the plan of Buglio, Forcola and Ardenno, allowing different changes in the riverbeds during to spring and summer flooding. In 1841 the project, supported by the Austrian government, was to force the Adda to flow into a new traced riverbed from the opposite side with respect to the old path, not far from the houses of Selvetta and Sirta. For this reason, it was necessary to dig a new riverbed more than 4 km long, to make new streams, to fix the outskirts of the nearby streams and the military road giving to the area the current structure (Fig. 11a-bottom).

---

<sup>6</sup> <http://api.eurogeoss-broker.eu/>

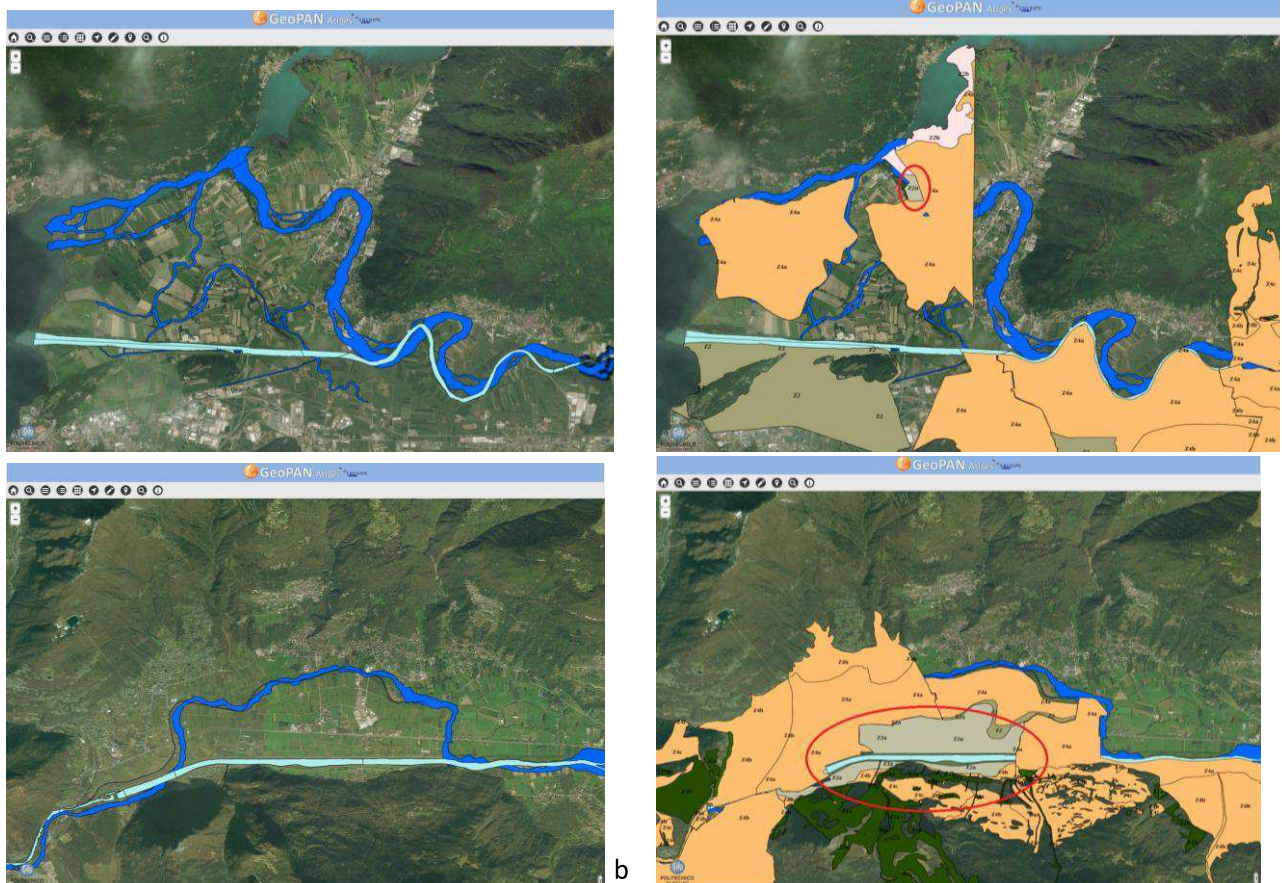


Fig. 11: Example of GeoPan application for the Adda river in the area of “Pian di Spagna” (top) and “Pian della selvetta” (bottom): (a) comparison between the ancient (blue) and the current (light blue) riverbed; (b) the link between riverbed and geological anomalies (areas classified as Z2a) is highlighted in red.

### 5.3 Zaragoza Histórica

As the GeoPan application, Zaragoza Histórica<sup>7</sup> aims at putting into value digitized and georeferenced historical cartography, but in this case taking advantage of the existence of a relatively substantial amount of large scale maps covering the evolution of a certain area (in this case, the city of Zaragoza, in the North-East of Spain) during a certain period of time (in our case, a four-century span).

Zaragoza Histórica is aimed at the general public and provides them with a visualisation tool to portray the historical cartography from the Zaragoza Municipal Archive presented in Table 1. Users of Zaragoza Histórica can select any of the 19 historical maps included in the collection and browse them individually by performing the usual actions of zooming in and out, panning, etc. Users are also able to compare any of them with the current cartography by overlapping the maps and setting a transparency level. Users are also provided with details of the selected map as year, author, dimensions, etc., and a link to the municipal archive web page is provided, where the complete metadata on the historical map is available. Additionally, tools common to other geographical visualisation applications provided by the Zaragoza Council are offered, like current toponym search.

<sup>7</sup> <http://idezar.zaragoza.es/visorHistorico/?locale=en>

Particular care was put into making the application extremely easy to use (Fig. 12), with the objective of reaching as many users as possible, and in-line with the Zaragoza Council policy of only requiring basic IT skills for using any of the web applications offered in their website. This tool is currently part of the offered services of the city SDI (IDEZar<sup>8</sup>) and makes use of previously deployed services and of the ENERGIC OD Virtual Hub.

The concept of Zaragoza Histórica can be easily replicated in other areas, provided the availability of several large-scale historical maps. From this perspective, the name “Zaragoza Histórica” can be considered the local brand under the application concept is operating in the case of Zaragoza.

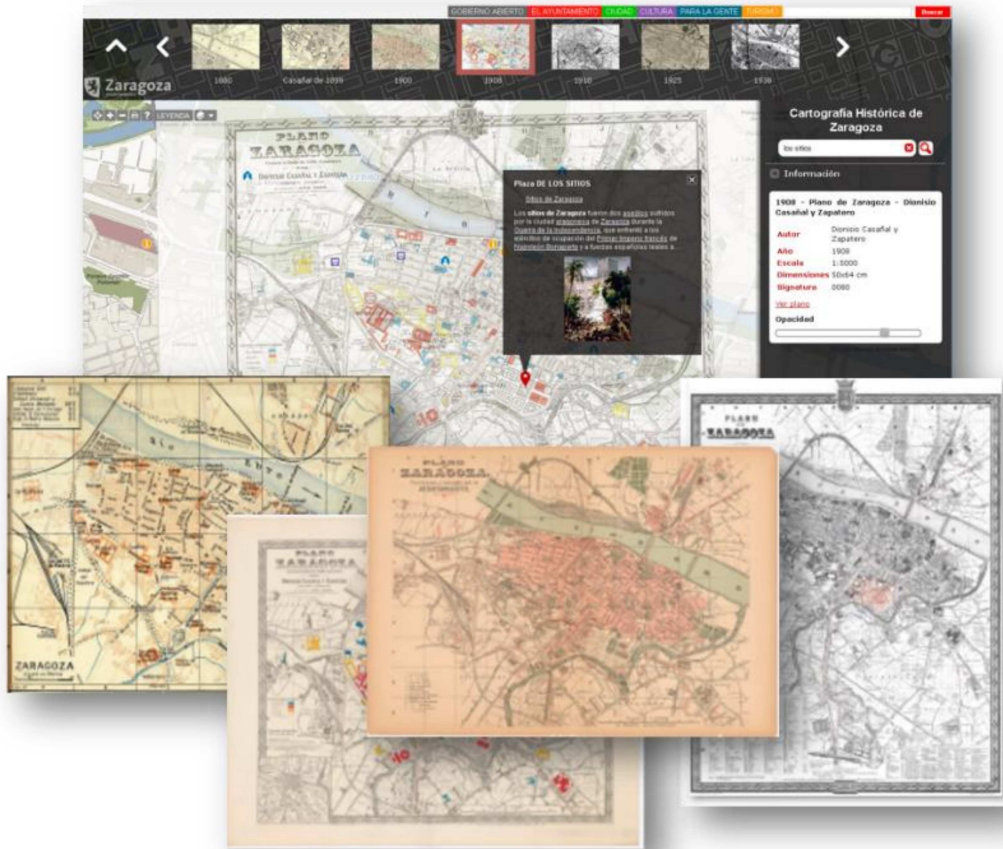


Fig. 12: User interface of Zaragoza Histórica and some examples of the historical maps of the city of Zaragoza included in it.

### 5.3.1 Development and software architecture of Zaragoza Histórica

In 2008, the Zaragoza Council published many historical documents coming from the Zaragoza Municipal Archive that had been previously digitised. Among them, several historical maps were included. Although the digitised historical maps were made available as Open Data, the format in which these maps were offered (raster DjVu files) limited to great extent their reuse and exploitation by the council itself and by third parties. DjVu (Bottou et al. 1998) is an open file format designed primarily to store digitised documents, very suitable for archiving and long term preservation of digital documents but it is not widespread and it does not allow for its contents (maps in this case) to be georeferenced.

<sup>8</sup> <https://www.zaragoza.es/ciudad/idezar>



1 The Zaragoza Council wanted to increase the reusability and visibility of these maps, while approaching  
2 them to the general public, so it was decided to offer these maps as an OGC service to (i) facilitate the  
3 reusability of these data and (ii) create a web application intended for the general public based in this OGC  
4 service. It was also thought that this represented a good opportunity for using ENERIGIC OD's Virtual Hub  
5 technology.  
6

7 The Council selected 19 of these historical maps which, as part of the ENERIGIC OD project, were  
8 transformed into a more common format (TIFF) and resampled to an appropriate resolution for the  
9 intended web application. These maps were also georeferenced and stored as GeoTIFF files in order to be  
10 used in the application, as explained in section 5.1.  
11  
12

13 Particular care was put also in the creation of metadata of these digital maps. ISO 19915 compliant  
14 metadata records were created, incorporating also some metadata about the original resources (the paper  
15 historical maps) that was available in MARC format. Future developments of the application can make use  
16 of these metadata records to allow the user to perform searches in a geographical catalogue and to  
17 automatically load into the web application services able to portray the results of these searches.  
18  
19

20 Regarding the application itself, the provision of the data related to historical maps through the chosen  
21 OGC interfaces was delegated into the Virtual Hubs in order to reduce the development effort. The VH is  
22 used to access the data through a Web Map Service (WMS) interface, while the VH itself access the original  
23 data (the historical maps GeoTIFF files) using the Web Accessible Folder (WAF) interface. Web Map Tile  
24 Service (WMTS) was initially the selected interface, due to its better performance in execution time.  
25 However, tests with WMS interface were conducted and the response times kept appropriate for an  
26 interactive application. Additionally, the use of the WMS implied less stress in terms of memory used by the  
27 VH, and eased the foreseen activities of adding new historical maps to the application, so finally the WMS  
28 interface was adopted. Zaragoza Histórica also access other data sources, like, for instance the *WMS Mapa*  
29 *Base Zaragoza* that provides the current cartography of the city, also through a WMS interface.  
30  
31  
32  
33  
34

35 The front-end of Zaragoza Histórica is composed of two different elements: the map viewer container and  
36 the viewer itself. The container, based in jQuery technology<sup>9</sup> (jQuery 1.8 + JQueryUI<sup>10</sup>), implements most of  
37 the interface offered to the user and interacts with the map viewer using the façade pattern. The second  
38 element, the map viewer, is implemented with a proprietary library developed by Universidad de Zaragoza  
39 and based in the Open Source framework OpenLayers<sup>11</sup> 2.12. It provides access to the base cartography  
40 and gives all the tools to navigate through the map. By default, it only offers a plain map and the  
41 functionality to draw different kinds of data over it in a dynamic way. This feature is used by the container  
42 to request and display all the special geographic data made available to the end user.  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57

---

58  
59 <sup>9</sup> <https://jquery.com/>

60 <sup>10</sup> <http://jqueryui.com/>

61 <sup>11</sup> <http://openlayers.org/>

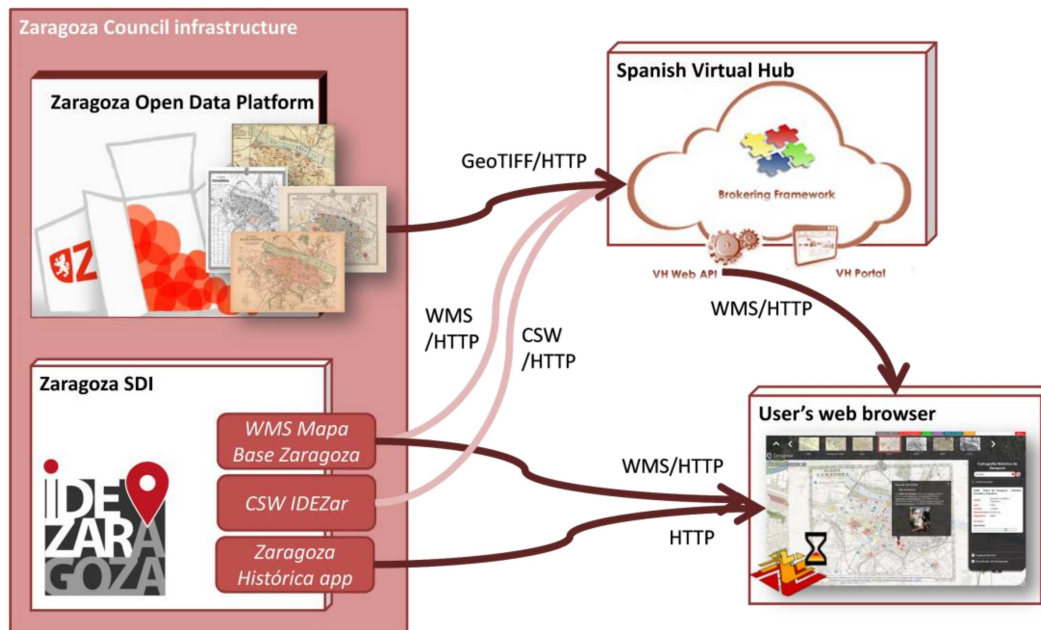


Fig. 13: Deployment diagram

Fig. 13 shows the deployment diagram of the application. Zaragoza Histórica runs in the user's web browser in his or her computer. The *Spanish VH* component runs in a set of three virtual machines considered as one single node conceptually. The historical maps in GeoTIFF format are stored and accessible in the *Zaragoza OD platform*, while the services from the *Zaragoza SDI* are located in a different node. Again, these last two nodes, which lay inside the *Zaragoza Council infrastructure* are indeed composed of several computers. This approach allows not only for the provision of the web application in the SDI portal, but also for the publication of the GeoTIFF files that are used in the OD platform and the OGC WMS service provided by the Virtual Hub. However, due to the size of GeoTIFF files, the Council of Zaragoza decided that only the WMS service provided by the VH is publicly available (the repository of GeoTIFF files weights 13.6 GB, while each request to the WMS for a 3 megapixel image in PNG format is on an average of 5 MB).

## 6. Conclusions

The many formats, interfaces, standards and protocols in used by different communities of users in the GI world represent an important barrier for an effective reuse and integration of geographical open data. In particular, software developers have to deal this complexity when developing applications that make use of GI, facing formats and standards they may be not familiar with.

In this paper, a brokered approach for the generation of web applications based on historical maps has been described. The presented approach uses a mediation layer, named Virtual Hub, that takes in charge of many of the technical issues concerned to interoperability between data, in particular spatial and geographic data, gathered from different data providers, with different features and referred to different historical periods.

The use of the brokered approach allowed reduced the effort needed for development of the both presented applications. Developers of GeoPan were able to avoid interact with geographical data directly by using the Virtual Hub Web APIs. In contrast, developers of Zaragoza Histórica used the Virtual Hub to access the historical maps though OGC services, but delegating in the VH the deployment of these services. It is important to highlight how the interaction between various data sources and cartographical products gives more value to the analysis of the territory and the possibility of enlarging the knowledge about it. The

1 capability to integrate different informative layers, both historical and modern, can be an important  
2 opportunity of development with application areas still largely unexplored. In a so complex environment, it  
3 is evident how the possibility of having an architecture with a single point of access ensuring an easy  
4 interoperability between different sources becomes crucial.  
5

## 7. References

- 6  
7  
8  
9 Affek, A., 2013. Georeferencing of historical maps using GIS, as exemplified by the Austrian Military Surveys  
10 of Galicia. *Geographia Polonica*, 86(4), 375-390.  
11  
12 Assante, M., Candela, L., Castelli, D., Pagano, P., 2014. The D4Science research-oriented social networking  
13 facilities. *ERCIM News*, 96, 44-45.  
14  
15 Assante, M., Candela, L., Castelli, D., Coro, G., Lelii, L., Pagano, P., 2016. Virtual research environments as-a-  
16 service by gCube. *PeerJ Preprints*, 4, e2511v1.  
17  
18 Balletti, C., 2006. Georeference in the analysis of the geometric content of early maps. *ePerimtron1* (1):  
19 32-42.  
20  
21 Barazzetti, L., Brumana, R., Oreni, D., Previtali, M., 2014. Historical Map Registration via Independent  
22 Model Adjustment with Affine Transformations. In *International Conference on Computational Science and  
23 Its Applications* (pp. 44-56). Springer, Cham.  
24  
25 Barron, C., Neis, P., Zipf, A., 2014. A comprehensive framework for intrinsic OpenStreetMap quality  
26 analysis. *Transactions in GIS*, 18(6), 877-895.  
27  
28 Bottou, L., Haffner, P., Howard, P.G., Simard, P., Bengio, Y., Le Cun, Y., 1998. High Quality Document Image  
29 Compression with DjVu. *Journal of Electronic Imaging*, 7(3), 410–425.  
30  
31 Béjar, R., Latre, M. Á., Nogueras-Iso, J., Muro-Medrano, P. R., Zarazaga-Soria, F. J., 2009. Systems of systems  
32 as a conceptual framework for spatial data infrastructures. *IJSDIR*, 4, 201-217.  
33  
34 Bigagli, L., Santoro, M., Mazzetti, P., Nativi, S., 2015. Architecture of a process broker for interoperable  
35 geospatial modeling on the Web. *ISPRS International Journal of Geo-Information*, 4(2), 647-660.  
36  
37 Boldrini, E., Papeschi, F., Santoro, M., Nativi, S., 2015. Enabling interoperability in Geoscience with GI-suite.  
38 In *EGU General Assembly Conference Abstracts* (Vol. 17).  
39  
40 Bonham-Carter, G. F., 2014. *Geographic information systems for geoscientists: modelling with GIS* (Vol. 13).  
41 Elsevier.  
42  
43 Brovelli, M. A., and Minghini, M., 2012. Georeferencing old maps: a polynomial-based approach for Como  
44 historical cadastres. *e-Perimtron*, 7(3), 97-110.  
45  
46 Candela, L., Pagano, P., 2015. Cross-disciplinary data sharing and reuse via gCube. In: *ERCIM News*, Issue  
47 100, January 2015 [http://ercim-news.ercim.eu/en100/special/cross-disciplinary-data-sharing-and-reuse-  
48 via-gcube](http://ercim-news.ercim.eu/en100/special/cross-disciplinary-data-sharing-and-reuse-via-gcube)  
49  
50 Candela, L., Castelli, D., Coro, G., Lelii, L., Mangiacrapa, F., Marioli, V., Pagano, P., 2015. An infrastructure-  
51 oriented approach for supporting biodiversity research. *Ecological Informatics*, 26, 162-172.  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65



- 1 Cerreta, M., Mele, R., 2012. A landscape complex values map: integration among soft values and hard  
2 values in a spatial decision support system. In International Conference on Computational Science and Its  
3 Applications (pp. 653-669). Springer, Berlin, Heidelberg.
- 4 Denzer, R., Schlobinski, S., Boot, G., Keppel, F., and de Rooij, E., 2014. An information platform fostering re-  
5 use of water data. In: International Congress on Environmental Modelling and Software (2014),  
6 <http://www.iemss.org/sites/iemss2014/proceedings.php> ISBN: 978-88-9035-744-2  
7  
8
- 9 Dihé, P., Denzer, R., Schlobinski, S., 2015. An Information Model for a Water Information Platform. In ISESS  
10 (pp. 91-101), *Infrastructures, Services and Applications. ISESS 2015. IFIP Advances in Information and*  
11 *Communication Technology*, vol 448. Springer.
- 12  
13
- 14 Fielding, R. T., 2000. Architectural styles and the design of network-based software architectures, University  
15 of California, Irvine, 2000.
- 16  
17
- 18 Fielding, R. T., Taylor, R. N., 2002. Principled design of the modern Web architecture. *ACM Transactions on*  
19 *Internet Technology (TOIT)*, 2(2), 115-150.
- 20
- 21 Grabowski, R. C., and Gurnell, A. M., 2016. Using historical data in fluvial geomorphology. *Tools in Fluvial*  
22 *Geomorphology*, 56-76.
- 23
- 24 Giuliani, G., Nativi, S., Obregon, A., Beniston, M., Lehmann, A., 2017. Spatially enabling the Global  
25 Framework for Climate Services: Reviewing geospatial solutions to efficiently share and integrate climate  
26 data & information. *Climate Services*.
- 27  
28
- 29 Goodchild, M., Egenhofer, M. J., Fegeas, R., Kottman, C. (Eds.), 2012. *Interoperating geographic information*  
30 *systems (Vol. 495)*. Springer Science & Business Media.
- 31  
32
- 33 Grosso, E. , 2010. Integration of historical geographic data into current geo-referenced frameworks: a user-  
34 centered approach. In *Proceedings of the 5th International Workshop on Digital Approaches in*  
35 *Cartographic Heritage (Vienna 2010)* (pp. 107-117).
- 36  
37
- 38 Guigoz, Y., 2016. *Spatial Data Infrastructures for addressing environmental challenges: stocktaking, capacity*  
39 *building, implementation and assessment (Doctoral dissertation, University of Geneva)*.
- 40
- 41 Haklay, M., 2010. How good is volunteered geographical information? A comparative study of  
42 OpenStreetMap and Ordnance Survey datasets. *Environment and planning B: Planning and design*, 37(4),  
43 682-703.
- 44  
45
- 46 Hell, T., Kohlhas, E., Schlobinski, S., Denzer, R., Güttler, R., 2013. An information system supporting WFD  
47 reporting. In *International Symposium on Environmental Software Systems* (pp. 403-413). Springer, Berlin,  
48 Heidelberg.
- 49
- 50 Jenny, B. (2010). New features in MapAnalyst, *e-Perimtron*, 5, pp. 176–180.
- 51
- 52 Lassoued, Y., 2012. D4. 4.2–Implementation of the Semantic Framework–Version 2. NETMAR Consortium.
- 53
- 54 Leadbetter, A. M., Lowry, R. K., Clements, D. O., 2014. Putting meaning into NETMAR–the open service  
55 network for marine environmental data. *International Journal of Digital Earth*, 7(10), 811-828.
- 56  
57
- 58 Mazzetti, P., Latre, M.Á., Bauer, M., Brumana, R., Brauman, S., Nativi, S., 2015. *ENERGIC-OD Virtual Hubs: A*  
59 *brokered architecture for facilitating Open Data sharing and use. eChallenges e-2015 Conference, Vilnius*  
60 *(Lithuania), 25-26 November 2015*.
- 61  
62  
63  
64  
65

- 1 Nativi, S., Craglia, M., Pearlman, J., 2012. The brokering approach for multidisciplinary interoperability: A  
2 position paper. *IJSDIR*, 7, 1-15.
- 3 Nativi, S., Craglia, M., Pearlman, J., 2013. Earth science infrastructures interoperability: the brokering  
4 approach. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(3), 1118-  
5 1129.
- 6  
7  
8 Nativi, S., Mazzetti, P., Craglia, M., Pirrone, N., 2014. The GEOSS solution for enabling data interoperability  
9 and integrative research. *Environmental Science and Pollution Research*, 21(6), 4177-4192.
- 10  
11 Ortiz Jr, S., 2007. Getting on board the enterprise service bus. *Computer*, 40(4).
- 12  
13 Podobnikar, T., 2009. Georeferencing and quality assessment of Josephine survey maps for the  
14 mountainous region in the Triglav National Park. *Acta Geodaetica et Geophysica Hungarica*, 44(1), 49-66.
- 15  
16 Regueiro, M. A., Viqueira, J. R., Taboada, J. A., Cotos, J. M., 2015. Virtual integration of sensor observation  
17 data. *Computers & Geosciences*, 81, 12-19.
- 18  
19  
20 Schmidt, M. T., Hutchison, B., Lambros, P., Phippen, R., 2005. The enterprise service bus: making service-  
21 oriented architecture real. *IBM Systems Journal*, 44(4), 781-797.
- 22  
23  
24 Shahand, S., Benabdelkader, A., Jaghoori, M. M., Mourabit, M. A., Huguet, J., Caan, M. W., ... and  
25 Olabarriaga, S. D., 2015. A data-centric neuroscience gateway: design, implementation, and experiences.  
26 *Concurrency and Computation: Practice and Experience*, 27(2), 489-506.
- 27  
28  
29 Southall, H., 2014. Applying historical GIS beyond the academy: four use cases for the Great Britain HGIS. In  
30 *Toward spatial humanities: historical GIS and spatial history*. Indiana University Press.
- 31  
32 Torres-Bejarano, F., Denzer, R., Ramírez, H., Hell, T., Frysinger, S., Schlobinski, S., Güttler, R., 2011.  
33 Development and integration of a numerical water quality model with the geospatial application suite CIDs.  
34 MODSIM 2011 - 19th International Congress on Modelling and Simulation - Sustaining Our Future:  
35 Understanding and Living with Uncertainty, pp. 1272-1279.
- 36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65