ArCEs A Digital Archive of Italian Colonial Cartography and Scientific Expeditions

Irene Calloud
Department of Historical and Geographical Studies, University of Florence, Italy
irene.calloud@unifi.it

Paola Zamperlin
Department of Historical and Geographical Studies, University of Florence, Italy
paola.zamperlin@unifi.it

Abstract

ArCEs is a study for developing a digital archive on Italian scientific expeditions in Northern and Eastern Africa and the former Italian colonies (19th–20th centuries). The aim of the project is to assess, protect and enhance an important corpus of documents (historical cartographies, photographs, scientific papers and archive documents) distributed among public archives and private collections. The database structure is based on the Dublin Core metadata standard. The information system is designed to integrate and make interoperable digital resources, to ensure standardized and complex indexing, and to support advanced retrieval, according to the standards in use. The geolocation of the resources in a GIS environment can display query results in the Google Earth environment.

Key Words: historical maps; colonial cartography; geo-referencing; webgis; geographic retrieval; open access
1. Introduction

The idea of the project is to preserve and make available on the web a corpus of archival documents on Italian explorations in the former colonies from the end of the 19th until the beginning of the 20th century. Using the internet, the materials are made available to a wide audience and additional features allow further map studies in ways that would not be possible on delicate and irreplaceable originals.

In recent years Italian researchers have shown renewed interest in the history of the colonial period and ever more often scholars need to access cartographic, photographic or archival sources concerning the Italian expansionist policy towards Libya, Eritrea, Somalia, Dodecanese and Albania. Unfortunately, documents are frequently not accessible or scattered among several archives.

Some recent projects and studies have well stressed the role of cartography in military politics during this historical period, and have investigated the connections between geography, cartography and military intelligence (Casti & Turco, 1998; Boria, 2007; Casti, 2007). As is well known, the Italian government funded geographical, but also archaeological expeditions with the goal of a political, urban and agrarian conquest (Carazzi, 1972; Cerreti, 2003; Ghezzi, 2003; Umiltà, Barani, & Bonati, 2004; Secciani, 2005; Spagnoli, 2010).

However, a great deal of research has often been limited due to the impossibility of analyzing a number of documents which were hard to obtain because they were kept in smaller, private or lesser known institutes. These preserve unique and precious documents testifying to the vast Italian cartographic activity that, for the first time, recorded geographical areas never surveyed before, except in a few English or German maps. Italian cartographers developed new technical methods of survey, solving problems in surveying areas which were completely different from Italian landscapes, such as the Sahara Desert. Moreover, accuracy was required for military needs, especially for toponyms used before the Italian conquest.

Believing that knowledge must be shared, the Applied Geography Laboratory of the University of Florence decided to design and create a digital archive for such documents, namely a collection of heterogeneous resources consisting of maps, photos, notebooks, papers and gazetteers regarding colonies in...
Northern and Eastern Africa, available only on paper, without homogeneous cataloguing and sometimes in a precarious state of conservation.

The project is designed to preserve, archive and publish on-line this historical and cultural heritage, cataloguing maps previously outside inventory.

Methods and techniques will be replicable on other documents, different from the present case study, because the system is based on standard methodologies of data collection. In addition, we are working to provide interoperability with other systems, meeting all the requirements needed. The final goal will be the integration of ArCEs into multidisciplinary projects that are standard compliant (Azzari, Calloud, & Zamperlin, 2011).

The main aim of this project is to provide an accessible and user friendly tool for a very wide audience — a place in which to find rare documents, mainly cartographic, thanks to the combination of internet and GIS. “Library science has experience in designing catalogues and, more recently, search engines that allow users to find the resources they require. Adding searching by place requires the use of locational information, either in the form of coordinates, or as place names, or both” (Gregory & Ell, 2007, p. 145). This requires combined efforts among researchers working in different fields (librarians, geographers, archivists, and IT experts) in order to make this experience valuable and to respond to the problem of finding relevant data, which is crucial when searching the Net.

We also must not forget that placing maps on the internet makes them available to a very vast public, while preserving the originals — often damaged or, at least, very delicate — and allowing researchers to elaborate on their studies, using the many tools available thanks to modern technology.

2. Fundings

The realization of this project was possible thanks to funding by the Region of Tuscany and structured European funds (POR CREO FESR 2007–2013 Activities 1.1 Intervention guidelines d). The overall duration is two years. The funding, although scarce when compared to the amount of study and work, covers a feasibility study and the consequent implementation of a prototype of a digital cartographic archive limited to the colonial exploration
theme, which can then be extended to the whole historical cartography preserved in the Geographical Library of the Florentine Athenaeum.

3. State of the art

During the last few years the most important European Libraries have faced the task of following developments in digital cartography. The majority of them are concerned simply with scanning, but in some cases they embrace new geographical retrieval technologies.

Digital libraries and internet can potentially make data available to researchers in ways that were unimaginable in the past. Nevertheless, actual data preservation and dissemination require competences and efforts that outstrip single research project objectives. Skills and specialized structured organization are fundamental, such as can be found in the Arts and Humanities Data Service (AHDS) and the UK Data Archive (UKDA) in the UK, and the Alexandria Digital Library and California Digital Library in the US (Gregory & Ell, 2007).

The Alexandria Digital Library (ADL) is a very interesting example. Conceived in the 1990s to solve the traditional access and organizational problems in map libraries, it offers researchers a wide repertory of georeferenced objects accessible through the Web (Goodchild, 2004).

Considering projects expressly concerning historical cartography, we should remember one of the most renowned on-line portals of old maps: the National Library of Scotland’s website. It has the largest selection of digital maps in the UK with over 44,000 accessible high-resolution images, within an on-line map library in the ‘Digital resources’ section of their institutional website. What is also relevant is that the website is a collaborative on-line project, where anyone with Internet access can help with the georeferencing of various historical maps from the library’s collection. The National Library of Scotland is one of the initial contributors to Old Maps Online (http://www.oldmapsonline.org), an English digital library trying to create a search portal which covers historical maps in multiple collections, with over 60,000 maps at the launch in February 2012. It plays the role of aggregator, allowing users to access at once several external files already digitized, but so far difficult to navigate (other partners are: the British Library, the American David Rumsey Collection at the Moravian Regional Library in Brno).
Three earlier projects provided access to multiple digital collections. The first one is the European DIGMAP (Discovering our Past World with Digitised Maps) specialized in historical materials, displaying about 5680 maps stored in The National Library of Portugal, the Royal Library of Belgium, the National Library of Italy in Florence and the National Library of Estonia. It pursued the purpose of becoming the main old map and related bibliography international information source and reference service. Another example is CartoMundi, developed by the Maison Méditerranéenne des Sciences de l’Homme, which consists of an open network of members and offers easy access to the treatment of copies of 8000 maps.

In these frameworks, the Italian panorama offers qualified projects in historical map digital preservation. In other words, it means mostly static maps in high resolution formats or with an interactive view by zooming, panning or hyperlinking to other information. Their main goal is to spread knowledge and to make our rich historical heritage available on the web.

Giving a brief overview, we can mention:

- *Società Geografica Italiana (Italian Geographical Society)*
- *Istituto Geografico Militare (IGM)*, which has published on the web a catalogue of historical maps and an interactive catalogue that allows the user to identify, through a topographical search or by using a search box, topographic maps, digital map data, geographic databases, geodetic elements, orthophotomaps, aerophotos, historical maps, historical photographs.
- *Geoweb*: the cartographical and graphical website of the Marciana National Library in Venice, an on-line catalogue with more than 29,000 digitalized documents.
- *Imago Tuscie*: is an on-line digital archive composed of more than 2000 documents preserved in the Siena and Grosseto Archives. It allows the visualization of map reproductions in high resolution, linked to their dossiers and grouped into several categories.
- *Archivio Capitolino*: The Roman Historical Archives preserves and publishes on-line the documentation produced by the Municipality of Rome, from the 12th to the 20th centuries, including maps and views of Rome and Lazio.
• **RAPu** (Rete Archivi Piani urbanistici): a digital archive concerning Italian urban planning from the unification of Italy to nowadays. It deals with maps, drawings, reports and various documents concerning planning processes.

Dynamic/geo-referenced maps:

• **Castore**: an open access catalogue focusing on Tuscany’s land registers of the 19th century. The maps are freely available on the internet, both as individual replicas of the original archive and as a mosaic of georeferenced maps, overlapped on modern maps in a WEB-GIS.

• **The historical drainage/land reclamation information system** (Sistema informativo delle bonificate storiche): a digital library of historical maps of Monsummano (Pistoia) that includes explanatory texts, thematic maps and raster images.

• **SIGEC**: the Central Institute for Cataloguing and Documentation is institutionally responsible for the General Information System for Cataloguing, the SIGEC, the national system for the acquisition and integrated management of knowledge of Italian cultural heritage. The cataloguing data are connected with GIS functions. It contains cross-referencing relations between the different typologies of heritage and displays relationships with the environmental and human elements.

4. **Project stages**

4.1. **Document selection**

Our case study analyzes three lesser known colonial archives, all stored in Florence.

The choice to focus our prototype on these three derives from the fact that these archives are homogeneous both in their chronology and geographical areas of interest, even though the preserved items differ. Two of these archives are inedited, and preserved in private collections. Thanks to the coherence of the core contents, these provide the perfect case upon which to create a pilot project. Starting from this initial project the intention is to subsequently create an exportable platform enriched with other documentary resources — especially cartographical ones.
The archives used are:

- **Marinelli’s Archive**: This repository has been preserved at Florence University’s Geographical Library since the 1930s. It was a donation from Giovanni Marinelli’s (Udine 1846-Florence 1900) heirs. He was a geographer and geography teacher at the Regio Istituto di studi superiori of Florence from 1893 to 1900. Olinto Marinelli (Udine 1876-Florence 1926), Giovanni’s son, was also a geographer and geography teacher in Florence from 1902 to 1926 (Patrizi, 2007). The nucleus of the analyzed maps in this project covers many colonial countries, including military maps of Libya (1913–1942), Dodecanese (1912–1943), Eritrea (1884–1941), Somalia (1890–1941) and Albany. Some of these maps were made by the Italian Military Geographic Institute (IGM), the official Government Cartographic Body. Although the archive does not have a real catalogue, a brief manuscript list is available to access the collection.

- **Bencini’s Archive**: Giovanni Bencini was a topographer and engineer employed in the Horn of Africa. His private collection preserves a considerable amount of black-and-white photos of some cartographic surveys carried out during the 1930s.

- **Caputo’s Archive**: Giacomo Caputo’s private collection; he was the last Italian conservator of Libyan Antiquities. The archive is stored in Florence and holds a great deal of unpublished documents on archaeological field surveys in Cyrenaica and Tripolitania (1935–1945), as well as photos, notebooks and papers (Calloud, 2012).

### 4.2. Digitalization process

After selection of the documents, they are digitized. Maps are scanned to produce a digital copy as raster data. This is a straightforward but tricky process, requiring a large and expensive scanner to manage big map sheets.

Many artefacts are in poor condition and are not usually available because of the extreme care with which they must be handled.

The maps have been digitized using scanners so as to allow the recognition and vectorization of features such as roads, settlements and rivers, which can be easily identified and singled out.
Obviously, the final image resolution can affect the accuracy of the resulting data. The scanning process needs a certain amount of post-processing. In fact, during the first digitization we adopted a resolution of 600 dots per inch, and afterwards the dimension of acquired maps was reduced to a web resolution (Catizzone & De Filippo, 2010, pp. 51–57).

4.3. Geo-referencing

The next stage of the project is the geo-referencing process (Hill, 2004). As is known, it is often impossible to georeference a cartographical representation if this does not have geometrical features that allow for the correct positioning without excessive deformation. Some bird’s eye view or demonstrative sketched charts that do not apply to a homogeneous reduction scale, cannot obviously be georeferenced like the geometrically precise maps. Nevertheless, it is possible to place these documents on georeferenced basemaps and to acquire important information from them (Azzari, 2010).

Historical maps are geo-referenced in a WGS84 spatial reference system, using ground control points identified on Google and Bing maps or using as reference points the intersections of the coordinate grid.

Concerning GIS procedures, it was decided to adopt both Esri/ArcGIS and Open Source (QGIS) technology, testing the latest software versions and the geoprocessing tools (Knowles, 2002; Gregory & Ell, 2007).

The non-cartographical objects in the database (pictures, sketches, articles, etc.) were all georeferenced as punctual, that is to say that they were all positioned in axes, so that it was possible to visualize them even through a satellite query.

4.4. Geodatabase and metadata

To ensure long-term usability, a database must be accompanied by suitable metadata, describing its contents, quality, conditions and other characteristics of the data held within a database.
Among the many existing metadata models, the ArCEs database is based on Dublin Core Metadata, adopted to provide interoperability for metadata vocabularies and semantic web implementations. In particular, we customized the Simple Dublin Core Metadata Element Set (DCMES) consisting of 15 metadata elements (Identifier, Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Source, Language, Relation, Coverage, Rights (Gregory & Ell, 2007, p. 57).

4.5. Design and development of WebGIS application

The University of Florence Applied Geography Lab has implemented a web-based application using dynamic maps and geo-referenced records. The application is designed to be reusable, updatable, and to conform to the main technology interoperability standards. The system can manage digital objects from different sources and different formats.

The scheme below represents the technical infrastructure of the ArCEs architecture and the figures 1–3 also show a back end system for managing and organizing documents.

*Fig. 1: Technical infrastructure of ArCEs architecture.*
The ArCEs WebGis provides spatial information to the digital object allowing its geographical localization and retrieval.

To achieve this, the project has been developed in J2EE and MVC standards. The software uses open source frameworks: Openlayers (map display), Hibernate (database management and data persistence); Spring. The WebGIS ArCEs is provided with a suitable tool for reading OGC WMS services, so the user can add geographical levels provided by other map servers. The GIS server used is Geoserver (Pridal & Zabicka, 2008).

The activities planned for this objective are divided into three main phases, still in progress:

1. Implementation of the digital repository
2. Data Entry
3. Testing and debugging

The ArCEs software is a WEBGIS exclusively developed with open-source libraries and adheres to the standards defined by the Open GIS Consortium.

The frameworks used for development are:

- GWT (Google Web Toolkit): a development toolkit for dynamic browser-based applications. Its great potentiality is the compatibility with most popular browsers (Internet Explorer, Mozilla Firefox, Google Chrome) and it gives the perception of a consultation very close to a desktop application.
- Hibernate: an open source middleware platform for developing Java applications that provides an ORM (Object Mapping Relational) service, managing the representation and maintenance of a Java objects database system. This framework guarantees platform independence from the database.
- Spring: its primary mission is to simplify the development of applications based on Java Enterprise Edition by restricting the configurations to be set and minimizing the dependencies between objects, providing a more robust application facilitating the reuse of components and making the development easier.
- PostgresSQL/Postgis: a complete ORDBMS whose characteristics compete with those of similar high level commercial products;
Fig. 2, 3: Back-end system for managing and organizing documents.
it is also equipped with a spatial cartridge for the management of geographic data.

- **Openlayers**: java.Script is a library for displaying maps and geographic data in web applications and is compatible with most browsers. With this library it is possible to develop web-based mapping applications similar to a desktop application. Openlayers is a project of the Open Source Geospatial Foundation.

- **Geoserver**: an open source server software written in Java allowing the publication and editing of spatial data. Designed to interoperate with other GIS systems, it is able to publish data that reside in the major sources of spatial data using open standards. Geoserver is the reference project for the implementation of standards like the OGC Web Feature Service (WFS) and Web Coverage Service (WCS); it is also a high performance WMS (Web Map Service). [GeoWebCache http://geowebcache.org/]

- **JBoss**: an open source Application Server that implements the full suite of Java EE. Being based on Java, the JBoss Application Server is multi-platform, usable on any operating system that supports Java. JBOSS has many components that ensure high reliability and robustness.

5. **Web interface**

When we designed the ArCEs system, we assumed that different users have different ways of accessing the same contents. In the last few years, Web 2.0 applications have shown that people want to manage, collect and share knowledge for their own purposes and they approach information with multiple strategies for searching. Sometimes users know what they want to find, but they usually ignore the existence of some documents or some information before searching; they are simply just searching.

However, library resources are often accessible only through traditional catalogues, in both digital format or not.

Clay Shirky claimed that “in the digital world there is no shelf, there is no physical constraint that’s forcing this kind of organization on us any longer”, so, according to him, we can realize that it is possible to get a resource without knowing its shelf number (Shirky, 2005). We need this number only when
we want to obtain the physical copy. This opens up to librarians countless opportunities to improve knowledge of the contents of their archives for as many users as possible.

In the case presented in this article, we are dealing with geographical information that is distributed in several sources, such as maps, photographs and aerial photographs, books, reports, archaeological surveys, notebooks with personal annotations, sketches, all not originally-digital documents and in different physical locations.

For these reasons, the ArCEs portal intends to offer various methods of providing access to historical maps and related documents. The Web interface has to focus attention on the user’s needs (Figure 4). We started by asking ourselves simple questions about which characteristics our potential users might have, and which needs they could or could not express in the searching process.

However, we are not librarians but geographers, so our approach may be a little different from that of librarians, because it is more interested in the content (forms, sizes, types and origins of geographical data) than in the object itself. In fact, libraries preserve historical documents that we can reuse or compare with other historical and current information, creating new knowledge. The availability of digital copies makes this process easier. Web interfaces and new geographic retrieval technologies can also facilitate access to contents, suggest to users multiple ways of reading historical sources in relation to the present and attract non-specialist users.

Fig. 4: Website design.
In designing the website for accessing our documents, we focused on three main solutions that represent three different access points:

1. a text list, by which it is possible to browse a traditional catalogue of all digitized materials (i.e.: search by title, subject, type of object, date, etc.);
2. a geographical search, for all document types, not only for maps, that transcends the traditional metaphor of searching;18
3. a relational search, based on the relationships between contents that we can visualize through graphs.

Figures 5 and 6 display some examples of geographical searching output, where we can see various historical objects located on a basemap (in this case a Google map). Using dynamic mapping applications we can increase the

Fig. 5, 6: Geographical searching output.
retrieval potential of traditional indices: through one visual application it is possible to reach information, by integrating different sources and presenting them in multiple layers and scales of zoomable mapping in a clear, intuitive way. Thus the searching process becomes a discovery process, because the query starts from the geographical area using co-ordinates and a specific period to find any unknown sources. We assume that each piece of information has a geographical component and so we can state that this geographical component could be the key to retrieving related information.

When you find a source, only then can you view the scanned and georeferenced image of the document and learn its physical location in a certain archive.

6. Conclusions

ArCEs provides easy access to documents that have been largely unknown until now and shares geo-cartographical data, both as an information resource and as a research subject.

The current principal projects include the on-line publication of the ArCEs website in December 2012; expanding the ArCEs database with new historical resources and geographical data; integrating the ArCEs system into multidisciplinary projects under the broad adoption of standards (i.e. SIGEC and other MiBAC initiatives); and, of course, fundraising.

References


**Notes**

1 The ongoing project is funded by the Regione Toscana (POR CREO FESR 2007–2013 Attività 1.1 Linee d’intervento d). The overall duration is two years. ArCEs leader: LabGeo. Applied Geography Lab, University of Florence. Scientific coordinator: Prof. Margherita Azzari, azzari@unifi.it. Partnership: Parallelo s.a.s (digital archive: project and application) and Geosystems S.r.l. (WebGIS).

2 www.geografia-applicata.it.

3 For a round-up of national and international digital projects concerning early mapping, see: http://maphistory.info/projects.html

4 http://geo.nls.uk/. The National Library of Scotland has one of the largest map libraries in the world, with around two million items. The Library’s collections span some 700 years, ranging from medieval manuscript maps to current digital mapping. Concerning the web portal see: Fleet, Withers Charles, and Wilkes (2011).

5 The Georeferencer application has been developed by Klokan Petr Pridal and the Moravian Library Brno as part of the OldMapsOnline project by the University of Portsmouth (UK), see http://www.oldmapsonline.org, Southall and Pridal (2012) and Fleet and Pridal (2012).
Also worth remembering is the Greifswald University’s cartographic collection, which holds about 7,000 historical maps of Denmark, Germany and Sweden: http://www.dhm.uni-greifswald.de/; and maps online of the Copenhagen University Library and the National Library of Denmark: http://www.kb.dk/en/nb/samling/ks/digifaks.html.

It consists of about 1,300 volumes of which 111 are rare, 558 geographical maps (of which many are single sheets), 213 containers, 57 brochure volumes, 320 various magazines. The geographical themed material is mostly of the 1800s and 1900s (Cassi, 1986).

It is an inedited archive, and its publication, thanks to Giovanni Bencini, will unveil the work of topographical engineers during the mapping of the colonial territories.

An interesting example of geographical search in library catalogues is in Buckland and Lancaster (2004).