A NEW INFORMATION SYSTEM FOR THE SUPERINTENDENCE OF POMPEII. INTEGRATION, MANAGEMENT AND PRESERVATION OF ARCHAEOLOGICAL DIGITAL RESOURCES IN THE PERSPECTIVE OF INTEROPERABILITY WITH EUROPEAN DIGITAL LIBRARIES

Rony Cesana
Sistemi Informativi S.r.l. – Liberologico
via F. Goya 2
56010 - Ghezzano (Pisa)
Italy
r.cesana@liberologico.com
http://www.liberologico.com

Maria Emilia Masci
Laboratorio LARTTE
Scuola Normale Superiore di Pisa
piazza dei Cavalieri 7
56126 – Pisa
Italy
e.masci@sns.it
http://www.sns.it

Andrea Vecchi
Sistemi Informativi S.r.l. – Liberologico
via F. Goya 2
56010 - Ghezzano (Pisa)
Italy
a.vecchi@liberologico.com
http://www.liberologico.com

Abstract – In April 2005 the Scuola Normale Superiore of Pisa started this Project aiming at the development of the Pompeii Information System. The project is financed by ARCUS SpA and developed together with the Archaeological Superintendence of Pompeii. The System has been designed by SNS and implemented by Liberologico Srl. The first step has been delivered in March 2008. The second step of the project will look into the design and publication of an official website and the installation of an OAI server in order to allow the harvesting of contents from national and European Digital Libraries and cultural Portals.

INTRODUCTION

The Project is funded by ARCUS S.p.a. [1] and developed by the Scuola Normale Superiore of Pisa (SNS) [2] together with the Archaeological Superintendence of Pompeii (SANP) [3]. It is regulated by two agreements among the three above-mentioned Bodies. The first one started in April 2005 and ended in March 2008. It produced the new SANP Information System. The System has been designed and developed by SNS in collaboration with Sistemi Informativi S.r.l. Liberologico [4].

The second agreement started in May 2007 and will end in October 2008. It will look into the design and publication of an official website and the installation of an OAI server to allow the harvesting of contents from national and European Digital Libraries.
and cultural Portals. It also will produce, in cooperation with the INDACO Department of the Politecnico of Milan, a prototype of 3D model representing the Forum of Pompeii, aimed at defining some best practices for data acquisition and rendering of 3D models.

The main requisite of the new SANP Information System is the ability to migrate various existing resources into one platform for the management of archaeological information related to the Pompeii site and the territory administrated by the former Superintendency of Pompeii, including Herulanum, Stabiae, Oplontis and Boscoreale. These resources include text, 2D and 3D images, audio and video documents, geographic information, and were previously managed by different software and stored in various repositories.

The System will become the main Information System of the SIAV (Archaeological Information System for the Geographic Area of Mount Vesuvius) and will be used both for internal purposes (cataloguing, documentation, preservation, management of archaeological heritage) and for external communication.

The unexpected creation of the new Archaeological Superintendency of Naples and Pompeii (SANP) in April 2008, which joins the two pre-existing Superintendencies of Pompeii and of Naples, brought into light the new need to integrate in the System also the digital resources pertaining to the Neapolitan area: an implementation plan is now under development.

Moreover, the System will permit the harvesting of its data and metadata from other repositories. Specifically, the data schemes adopted for Pompeii’s resources, which comply with the standard models of Italian Central Institute for Cataloguing (ICCD) [5], have been mapped into the Dublin Core Application Profile designed by SNS for CulturalItalia [6] under commission of the Italian Ministry of Culture. This mapping will allow the harvesting of the System’s contents from the Portal CulturalItalia [7] and interoperability with other repositories adopting the Simple and Qualified Dublin Core metadata descriptive standard [8].

**ANALYSIS AND REQUIREMENTS OF THE SYSTEM**

In 2005, when the Project started, the main objective identified by the three partners involved in the agreement was to collect in one System the existing digital contents that already were in possession of the SIAV.

Such contents pertained to the archaeological and cultural heritage administrated by the former Superintendency of Pompeii, that was extended in the large area around Mount Vesuvius.

**Main Objectives**

The analysis identified the following aims for the SANP System:

- to simplify the management, preservation and exploitation of the heritage administrated by the Superintendence with a System for the internal storage of data produced by its offices;
- to promote the research and exploitation of the Vesuvian patrimony amongst Bodies and people external to the Superintendence, by providing them with a System for data storage and interrogation flexible enough to manage information of different kinds, types, formats and structure;
- to disseminate high quality information to different kinds of users through a website and other media and devices.
Identification of the final users

Consequently, the final users of the SANP Information System have been grouped in the following types:

- **internal users**: all the users included in the SANP staff. They access the System through an authentication process. They are grouped into different groups depending on their office or activity, and can be granted different rights (consultation, data entry, data validation, system administration, etc.);
- **external users**: specialists working for Bodies external to the SANP (they can be archaeologists, architects, restorers, etc.), employed in projects and activities regarding the Superintendence or its heritage. They access the System through an authentication process. They can be grouped into different groups depending on their affiliation or activity and can be granted different rights;
- **indirect users**: this category joins different kinds of users that can’t directly access the System but can consult some of its contents through the website or through other media and devices.

Domain: contents to be managed by the System

The analysis identified very different kinds of existing digital resources to be integrated in the System, that were managed by different software and stored in various repositories. In particular, in 2005 the following data and repositories were stored in the SIAV office of the Superintendence:

- **Geographic data**: three Geographic Information Systems (GIS), managing various information:
  1. “Neapolis” GIS: themes related to the date of the excavated structures; index of the various kinds of architectural structures; information on the usage of buildings. Link to the databases. Cartography: the whole Vesuvian area, with a detailed cartography of the Pompeii site;
  2. “A plan for Pompeii” GIS: data about findings situated at, or originating from, the archaeological buildings; information on the conditions of the buildings, on the restorations and on the priority of preservation interventions. Cartography: Pompeii area;

- **Cataloguing data**: resources resulting from various cataloguing campaigns conducted by the Superintendence. All those information were managed by various software. They were not related to each others and were not interoperable. They mainly consist in:
  1. cataloguing charts of the archaeological objects (mostly paintings and mosaics from Pompeii) and monuments (private houses and public buildings from all over the Vesuvian area). These charts are structured according to the standard cataloguing schemes of ICCD or with schemes specially designed by the Superintendence for describing its peculiar heritage;
  2. digitised images (mainly drawings and photographs);
  3. charts and related digitised texts pertaining to the excavation diaries and reports by the archaeologists.
Along with these already existing contents, the analysis brought into light the opportunity to integrate in the System a huge quantity of information produced during the years by external actors. In fact, an effective exploitation strategy carried out by the Superintendence led to several agreements with Italian and foreign institutions and enterprises, for promoting and allowing researches, excavation campaigns, restoration activities, etc. Such external collaborations produced publications, exhibitions, databases, 3D reconstructions, documentaries and other high quality multimedia resources, that in the future could be stored in the System.

As a first step, an “Archive of the research projects” has been planned, including in the System special charts for describing this kind of activities and recording in the charts the results of a census of the projects carried out by external Bodies in accord with the Superintendence. An important objective for the future is to store in the System all the available digital products of the research projects.

In conclusion, the SANP Information System is designed to collect in one interoperable System the existing digital resources maintained by the SIAV and other resources on the Vesuvian heritage produced by external Bodies, also foreseeing the integration of other resources that will possibly be produced in the future.

On the basis of such results, the analysis identified as the main requisites of the System interoperability, scalability and flexibility.

**Requirements of the System**

The requirements of the System have been specified as follows:

- scalability and modularity of the System architecture: this requirement guarantees the durability of the System, by allowing the substitution of single modules without changing the whole System;
- possibility of supporting various formats: this requirement guarantees the ability of the System to integrate various kinds of resources;
- ability of representing complex relations between resources: this requirement guarantees the possibility of creating connections among different data;
- adoption of national and international standards: this requirement guarantees the interoperability of resources into the System and with external systems;
- use of open source software: this requirement guarantees the management and maintenance of the System during the time;
- development of a web-based System: this requirement guarantees the implementation of data into the System and their consultation by users dislocated in various geographic places. This requirement also implies the development of an appropriate users’ rights management and of an efficient workflow.

**SYSTEM ARCHITECTURE**

In order to satisfy the requirements on the adoption of standards, on the use of open source software and on the development of a web-based application, the SANP Information System exposes an XHTML interface compliant to the W3C’s WCAG standard and has been developed using JAVA technology: Google Web Toolkit has been adopted to simplify the deployment of AJAX features [9].

According to the scalability requirement, the System architecture has been designed as a composition of different modules for specific features: data managing, query
functionalities, XML import and export of data, repository management, data presentation and communication.

The Apache Tomcat servlet container and the MySQL DBMS have been used during development and testing stages, although virtually any J2EE container and any SQL-compliant DBMS could be employed.

The main system is made of three main modules:

1) **AxCore**: the core of the system. It is the JAVA library that offers an interface to the application layer client (web interface). It contains the modules for managing users’ rights, cataloguing charts, data, thesauri, vocabularies and authority files. Its features are accessible through five main interfaces:
   1.1) **AdminCatalog**: manages users, users’ rights (roles) and users working groups (projects);
   1.2) **MetadataRepository**: manages metadata schemas for the cataloguing charts and authority files;
   1.3) **DataRepository**: manages data pertaining to the cataloguing charts;
   1.4) **ThesMetadataRepository**: manages thesauri and controlled vocabularies stored in the system;
   1.5) **ThesRepository**: manages data pertaining to thesauri and controlled vocabularies.

2) **AxDbImpl**: implements all of the features published by the AxCore library. This implementation is based on a relational database.

3) **AxWeb**: it’s a web application that can be accessed through a web browser with support for Web 2.0 technologies. It uses the services exposed by AxCore.

The core functionalities are also available through a command line interface, published by a fourth module:

4) **AxTools**: is a JAVA library containing command line tools for the maintenance of the system. It contains:
   4.1) **AxReg**: for registering metadata schemas;
   4.2) **SelmoImport, XmlImport**: for importing XML data;
   4.3) **XmlExport**: to export a project’s charts onto an XML file;
   4.4) **TrcImport, TrcExport**: allow for import and export of charts, stored in files according to the ICCD standard file format (.Trc);
   4.5) **VVReg, VVImp**: allow for registration and data import of thesauri.

Geographic information is managed through the open-source server GeoServer, certified by the Open Geospatial Consortium (OGC). The web application exposing the map data uses the open source OpenLayers JavaScript library [10].

Finally, the OAICat open-source software have been added in order to integrate in the System a repository framework allowing the metadata harvesting using the Open Archive Initiative – Protocol for Metadata Harvesting (OAI-PMH) [11].

**SUPPORTED CATALOGUING SCHEMAS**

The system have been designed for the storage of contents structured according to the standard cataloguing schemas of the ICCD [5], which includes specific schemas for:

• **cataloguing charts**: for describing archaeological objects, monuments, sites;
• **authority files**: for describing bibliography and authors related to the catalogued heritage;

• **multimedia entities charts**: for describing and uploading in the System different kinds of documents (text, images, audio and video resources) in different formats, connected to the catalogued heritage.

To respect the requisite of flexibility of the System, it has been projected in order to manage new versions of the ICCD schemas and personalised schemas.

Thus, an additional tool (SchemaEdit) has been developed in order to author the definition of new cataloguing schemas; these can be then registered, therefore allowing for new types of data to be stored and be related to the existing contents. This feature allows to change and update the metadata schemas in the future.

For example, a specific metadata schema has been defined for the Project cataloguing chart, in order to start a census of the research projects carried out on the Vesuvian heritage during the last ten years. These charts can be related to each other and to the existing charts on the archaeological objects and monuments; also, they can be linked to multimedia entities with the related files attached.

All charts stored in the System can contain geographic references (point, poly-line and/or area) and, through such data, can be linked to the GIS, where all kinds of Shapefiles can be updated.

**MAPPING INTO NATIONAL AND INTERNATIONAL METADATA STANDARDS**

As one of the main tasks of the Project is to guarantee interoperability with other national and international repositories, an OAI repository has been integrated in the System to allow the metadata harvesting using OAI-PMH [11].

This solution will permit the harvesting of contents by the Italian Culture Portal “CulturaItalia” [7] and by other systems, after the mapping of the adopted schemas into Simple DC [8] and into the PICO Application Profile [6], specially developed for CulturaItalia, which further refines Qualified DC.

At the moment the ICCD cataloguing schemas used in the System for the description of archaeological objects and for the related bibliography have been mapped both into the PICO AP and into the DC metadata schemas. The related metadata configured according to the abovementioned metadata schemas have been exposed through the OAI repository, and harvesting tests have been successfully performed.

In this scenario, the SANP Information System and the related website, which is under construction, will provide metadata and contents to CulturaItalia and to other portals and digital libraries supporting Simple or Qualified DC, thus representing a model for Information Systems and repositories to be carried out by other public Bodies subjected to the Italian Ministry of Culture.

**MAIN FEATURES AND INTERFACE OF THE SYSTEM**

**Authentication**

Users access the System through an authentication process: after the log-in, they select one “Project”. Each “Project” defines a working group and a related activity.

Each user can access more than one Project with different “levels” or roles (e.g. one user can have the only-read right in one project and the data-entry right in another project). Every user can access the main project with at least the read-only level.
The main project, named “Catalogue”, is the general catalogue of the SANP joining all the cataloguing charts of all the Projects when they are in the status “approved”. After the log-in process, the current Project can be easily changed at any stage.

The main menu gives access to three main areas: Administration, Search and Browse, Data Entry.

Administration Area

The Administration area can be accessed only by users owning the right of “administrator”. The sub-menu of this area presents the following pages:

- **Projects administration**: Projects (working group/activity) can be created, modified and deleted here.
- **Users administration**: users can be created, modified and deleted here. All users are associated with their user names, passwords and personal data.
- **Users/Projects connection**: users can be assigned access roles for existing Projects (working groups/activities) here, and be granted set of rights.
- **Levels administration**: Levels can be managed here. A “Level” joins into one profile one or more users’ rights (e.g. a “Cataloguer” level would be associated with the following rights: read, write, link charts to the GIS).
- **Import of Charts**: here the administrator can import in the System one, or a group of, cataloguing charts in XML format or in Trc (ICCD format).
- **GIS administration**: here the administrator can upload in the system one or more Shapefiles.

Search and Browse Area

In this area all users can search, browse and view the contents of the Catalogue and of all the Projects which have been made accessible to them, with at least the read-only level. The sub-menu is structured as follows:

- **Simple search**: a Google-like search (free text search).
- **Advanced search**: permits to perform transversal searches through the different cataloguing charts schemas, by entering words or sentences in one or more search fields corresponding to Qualified DC terms.
- **Search by field**: allows performing an advanced search among charts with the same metadata schema.
- **Indexes**: displays indexes for each kind of cataloguing chart, authority file and multimedia entity stored in the system.
- **GIS**: gives access to the GIS interface.

After performing every kind of search or selecting an index, the interface displays a tree-shaped list in which the resulting types of cataloguing charts, authority files and multimedia entities are listed, flanked by the number of available results.

After selecting a type of content, user will access a list displaying the abstracts of the resulting charts. Each abstract is flanked by a drop-down list that allows the user to perform the operations for which he owns the rights. Every user will see the options:

- **Navigate**: displays the selected chart and its links with other charts and with the GIS, allowing the user to navigate through the contents of the system.
- **View**: displays the selected record in a PopUp-like window that can be closed.
- **Print**: displays a printable version of the chart.

Users with appropriate access level will be made available the following operations:

- **Modify**: for accessing the chart in the data-entry mode and modifying it.
• **Create a copy**: for creating a copy of the chart, after specifying an identification code for the chart to be created.

Only the users possessing the “administrator” right can select from the list of results one or more items, and press the button “Export”, to export them in XML or Trc format.

The GIS interface displays the map area and a tree-shaped menu on the left. Users can select one or more theme from the menu. They will be immediately displayed. By clicking on the icons in the visualization area, users can: zoom-in and out the map; perform panning (move into the map); select an area. After selecting an area into the map, the interface displays the list of the resulting charts related to that geographic area.

**Data Entry Area**

This area can be accessed by all the users who possess the “data-entry” right for the previously selected Project. The sub-menu displays the options:

• **Data Entry**: the cataloguer selects one type of chart, authority file or multimedia entity and starts entering the data.

The data-entry activity is aided by controlled vocabularies and authority files lists. Moreover, it is possible to look up all the values a specific field was filled with, select a value from this list to be used in the current chart, or to perform a “find and replace” operation with a given value into the current chart or in a group of charts.

A predefined workflow permits to define the “State” of the chart (“in course of implementation”, “to be approved”, “approved”, etc.). An integrated tool permits the formal and semantic validation of the chart in every moment.

The system automatically stores the previous versions of each chart.

• **User Manual**: displays the user manual.

**FUTURE DEVELOPMENTS OF THE PROJECT**

As previously specified, contents stored in the SANP Information System will be accessible for specialised users (“internal” and “external” users) after authentication. Some of the contents stored in the System will be also displayed for “indirect” users: this category joins different kinds of users that can not directly access the System, but can consult its contents through the web or through other media and devices.

Contents will be accessible on the web on an official website, which is currently under construction, or on national and international portals and digital libraries, that will harvest metadata via OAI-PMH.

The website of the Project will be conceived as a “Portal of the Vesuvian Archaeology”. The architecture of this Portal integrates four different areas:

• **Catalogue**: in this area it will be possible to consult contents deriving from the SANP Information System. They comprehend inter-related cataloguing charts, authority files and multimedia entities. All contents can be related to the GIS. Cataloguing contents will be displayed in a user-friendly visualization: charts will be simplified, avoiding the complicate partition in fields and subfields, and their data will be displayed in a more readable and accessible format.

• **2D and 3D GIS**: in this area geographic contents will be displayed and linked with the other resources available into the Portal. For the visualization of geographic contents Google Maps will be used: this solution seems to be more suitable and understandable for web users. By clicking on the map, it
will be possible to access: cataloguing resources; digital library resources; CMS contents; 3D models. 3D models will be used as a 3-dimensional GIS: some measurable and geo-referenced models will be accessible by clicking an icon in the map and displayed through a 3D viewer.

- **CMS contents**: a Content Management System will allow for the creation of specific web-pages that will be maintained by various Bodies and Institutions conducting their researches on the archaeological heritage of the Vesuvian Area. The Superintendence itself will use CMS to publish news and articles.

- **Digital Library**: edited and unedited texts, images and multimedia resources on the Vesuvian Archaeology will be stored in a Digital Library. Also materials published elsewhere on the web will be referenced here. The analysis identified the open source operating system Fedora [12] as the most suitable for this purpose.

In order to make available those contents through other media, a mobile version of the website will be developed.

Moreover, the INDACO Department of the Politecnico of Milan is working at a 3D model reproducing the Forum of Pompeii, that will serve as a prototype on the basis of which best practices and requisites both for data acquisition and for the rendering will be defined. CINECA research institute of Bologna will provide the Project with the Open Visman software, that will manage 3D models visualization both for stand-alone and for web applications.

References

[1] ARCUS S.p.a. is a company that promotes and develops arts, culture and performances, whose capital is entirely subscribed by the Italian Ministry of Economy and whose actions are planned by the Ministries of Culture and of Transport and Infrastructures. Website: http://www.arcusonline.org/.

[2] The Scuola Normale Superiore of Pisa is a university centre for teaching and research, founded by Napoleon in 1810 as a branch of the Ecole Normale Superieure de Paris. The LARTTE laboratory of SNS develops projects for the analysis, research, preservation, technology and economy in the field of Cultural Heritage. SNS Website: http://www.sns.it.

[3] In April 2008 the former Superintendence of Pompeii (website: http://www.pompeisites.org/) and the Superintendence of Naples and Caserta (website: http://www.archeona.arti.beniculturali.it/) have been joined by the Italian Ministry of Culture in one new Body: the Archaeological Superintendence of Naples and Pompeii (SANP). It comprehends the whole Vesuvian Area, the Neapolitan area (with the Archaeological Museum of Naples) and the area of Campi Flegrei.


[6] The PICO Application Profile is published at: http://purl.org/pico/picoap1.0.xml. For the mapping of ICCD schemes into the PICO AP see: Mapping dei metadati tra le schede ICCD e PICO Application Profile per il Portale della Cultura.

Dublin Core Metadata Initiative (DCMI) website: http://dublincore.org/.

For Google Web Toolkit, see: http://code.google.com/webtoolkit/.


Fedora Project website: http://fedoraproject.org/.