GIS And Web-GIS, Commercial and Open Source Platforms: General Rules for Cultural Heritage Documentation

Original

Availability:
This version is available at: 11583/2373419 since:

Publisher:
ISRPS

Published
DOI:

Terms of use:
openAccess
This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)
GIS AND WEB-GIS, COMMERCIAL AND OPEN SOURCE PLATFORMS: GENERAL RULES FOR CULTURAL HERITAGE DOCUMENTATION

F. Rinaudo¹, E. Agosto¹, P. Ardissone²

¹ Politecnico di Torino – DITAG – Torino – Italy
fulvio.rinaudo@polito.it, eros.agosto@polito.it, paolo.ardissone@polito.it

KEY WORDS: Spatial Information Sciences, Cultural Heritage, GIS, Standards, Open Source

ABSTRACT:
Cultural heritage documentation is the more generic term used to define the intelligent collection of all kinds of information needed to know an object in order to document it before restoration, conservation and management or just to share knowledge and to transfer it to the future.

The numerical formats used today to record all the information (from historical to the metric and/or physical data) requires the use of GIS technology to record and manage it.

The paper describes the analysis and tests performed by the authors concerning the potential of GIS and WEB-GIS on both commercial and Open Source platforms, and shows comparisons between the different possibilities of the GIS technology and a proposal for a possible common platform for Cultural Heritage documentation.

1. INTRODUCTION

After many years of separate and different activities directed towards surveying Cultural Heritage (CH) objects (man made or natural) from different points of view (e.g. metric survey, history, archaeology, anthropology, botany, construction methods, physical and chemical effects, etc.), the efforts of the international community are today devoted to promoting a more general approach.

All international organisations now use the term “Cultural Heritage Documentation” which indicates an interdisciplinary work where all the specialists involved in detailed investigations on an object of interest collect, interpret and share the data and the results of their interpretations on a common platform in order to allow a general understanding of the object itself and an integration of the collected information.

This result is due to the firm belief that only a close connection between different form of information can allow a complete understanding of the true significance and to plan future restoration, conservation and management activities.

Starting from the last decade of the last century, many researches on the use of GIS technology have been proposed to fulfill the goal of merging data originating from different sources. The basic idea was that only a common spatial reference system could allow a complete and deep integration, as has been demonstrated as far as land management is concerned.

All the researches and tests performed worldwide also showed the practical impossibility of an approach based on the experiences and standards proposed for Geographical Information Systems by CEN and ISO.

The different nature and contents of the investigations performed on a CH object need a more general approach, but it must be developed in an environment that is able to share information with GIS devoted to land planning.

Apart from these conceptual results a huge technical improvement in Information Technology (IT), has made new instruments (e.g. WEB-GIS, Open Source platforms, DBMS with geometric data) available which enable a possible more general standard approach in CH documentation to be outlined by using GIS (or rather Spatial Information Systems - SIS) technology. The general rules for a correct SIS for CH documentation and comparisons between the different technical solutions (GIS, WEB-GIS, commercial and Open Source platforms) are described in the next sections, concerning the results of the authors’ research activities.

2. GIS OR SIS FOR CULTURAL HERITAGE DOCUMENTATION?

The term GIS has been used mainly in Land management applications. The term “geographic” is traditionally connected to an international cartographic reference system. Considering the geometric deformation of the geographic reference systems, the term “geographic” can be misinterpreted, therefore it is better to talk about Spatial Information Systems (SIS) for Cultural Heritage applications of GIS technology. This assumption means that usually local reference systems (both continuous or discrete ones) are used to locate the acquired data; geographic coordinates (or global reference systems) can be used as ancillary information in order to locate the objects in a more general context.

SIS is a virtual space where each kind of information can refer a specific point in a known spatial reference system. Apart from this location of the data, the data themselves are connected in a logical environment (the Database) in order to allow an “intelligent” reading of the data.

Since 1990, SIS has been considered as an ideal instrument for the management of knowledge on Cultural Heritage. Starting form that period, all the organisations involved in the preservation and management of cultural heritage accepted SIS as the natural evolution of the more traditional databases.

The four points which justify this trend are:
- the significance of the spatial component of the collected data;
- easy accessibility to SIS technology;
- cultural heritage importance in the land planning strategies;
3. RULES FOR SIS IMPLEMENTATION

SIS implementation of a Cultural Heritage object is not an easy task. It can be stated that no general solutions can be offered in terms of logical data structure or investigation tools. Cultural Heritage Documentation and Management greatly depends on the author’s culture, the specific goals of the intervention, the epoch of the object, etc.

The correct implementation of a SIS for Cultural Heritage Documentation and Management can only be performed by a multidisciplinary team: the project team. The project team is formed by an SIS manager and a person in charge of each kind of investigation performed or to be performed on the investigated object. Each person in charge has to coordinate the team of the specialists in only one field of investigation and has to report the decisions to the project team. Each person in charge has to be able to accept or reject the proposals of the SIS manager in order to preserve the integrity of the inquiry to be performed.

Today, most SIS implementations require the recovering of old data and studies performed in the past by different specialists usually checked, integrated and/or completed by other different specialists.

All the involved specialists have to participate in the comprehension of the data and their interpretation. In the case where new data has to be collected the specialists have to discuss, with the SIS manager, how to organise the collection of the data considering not only the goals of the investigation but also the practical rules which allow a quick and protected input of the data inside the systems.

Each documentation specialist uses his/her own criteria to collect data: tables, symbols, key words, etc. Each of these criteria has to be discussed beforehand with the SIS manager.

3.1 Keywords list

Different aspects of an object can be recorded by different specialists. Therefore a common key words list, or at least an intermediate translation list, must be defined. As an example, if specialist A uses the term “date” to record a specific temporal event and specialist B uses the term “epoch” to record the same information, the possible solutions are: A and B decide to use the same attribute, or A and B use two different terms but these terms are grouped in the same attribute inside the database (e.g. the intermediate list says that “date” and “epoch” signify the same attribute of the element). This is not an easy goal to obtain. Specialists usually do not like to change a consolidated tradition. A possible solution could be the use of some accepted Thesaurus (e.g. British Museum Materials Thesaurus, Art and Architecture Thesaurus by Getty).

3.2 Data management common criteria

Many different researchers are usually involved in an investigation. Therefore common rules to acquire the information have to be drawn up between specialists in the same field. Common rules for bibliographic lists, date recording, image acquisition, etc have to be decided on before starting with the data collection.

3.3 Data quality parameter

A multidisciplinary environment and the international standards require that each piece of information has to be evaluated in terms of accuracy, completeness and consistency or, in a one word, quality. A general attribute (e.g. a date, an author, etc.) can be correct with different degrees of confidence. Quality estimation allows different values to be made of the same attribute and a critical and truthful reading of the collected information.

4. COMMERCIAL AND OPEN SOURCE PLATFORMS

If Internet is queried using the keywords “GIS” the number of answers in commercial solutions is similar to the Open Source environment.

Commercial platforms (e.g. ESRI, INTERGRAPH, etc.) were developed especially for Land planning but, thanks to the developing kits normally distributed to the final users, it is possible to customize the standard solutions into particular solutions that are able to solve all kind of problems.

Open Source resources are becoming increasingly more interesting especially for CH documentation. The two main reasons are:

- Cultural Heritage is not a private property, therefore its knowledge should be diffused and shared free of charge;
- Cultural Heritage documentation is not a standardized procedure but each object requires an ad-hoc solution.

The Open Source community now offers a complete set of instruments that are able to build up a complete SIS project (relational databases, graphic environment, geographical data management systems, etc.).

A comparison has therefore been made of the two approaches: commercial and Open Source, in order to choose the best opportunity.

The comparison was conducted by testing the following criteria:

- license costs
- maintenance costs
- easy development of the basic utilities
- easy development of advanced utilities
- customisation
- flexibility
- updating
- Web migrations

4.1 Licence costs

If one accepts that standard commercial solutions always need a customisation for Cultural Heritage applications, the license cost parameter stands out for Open Source solutions. In this case, both desktop platforms (e.g. GRASS), and relational databases (e.g. MySql, PostgreSQL) can be found free of charge.

4.2 Maintenance costs

The maintenance costs of Open Source Solutions are also favourable. All Open Source upgrades are in fact delivered free of charge, while in commercial solutions the users are forced to pay maintenance costs every year (of about 20-30% of the licence costs). Both in Open Source and commercial solutions, upgrades can require an upgrade of the customisations.
4.3 Easy development of basic utilities

Obviously, from this point of view, commercial solutions offer the best performance due to the high level of the basic functionalities offered by the standard versions of the tools. Open Source solutions offer the utilities developed by the user community but a detailed test of them and, in most cases, the building up of new solutions have to be taken into account.

4.4 Easy development of advanced utilities

A standoff can be applied for this evaluation element. In all cases, a skilled user is needed. Commercial solutions offer a limited way to setting up advanced solutions while the Open Source environment allows the use of different possible ways (e.g. with MapServer it is possible to use different scripting languages or the Java environment). Open Source pays a great deal of attention to the development of solutions that comply with international standards.

4.5 Customisations

Commercial platforms today allow good level of customisation using well known programming languages (e.g. Visual Basic for ESRI). However the only thing one can expect is that many of the offered tools are not useful for CH applications. The Open Source environment allow a free customisation and offers a wide community with free source scripts and the possibility of contacting the authors of the customisations in order to solve particular problems.

4.6 Flexibility

This property means the ability of the software to run on different hardware and OS solutions. From this point of view Opens Source solutions seem to be less constraining: commercial solutions allow easy running in the proprietary environment.

4.7 Updating

This term means the normal evolution of the software and the correction of the malfunctioning of the software itself. In this case, Open Source offers a more flexible and encouraging solution.

4.8 WEB migration

The GIS solution is usually understood as a desktop solution. Many books and manuals on GIS technology talk about a GIS application and its subsequent publication on Internet. Nowadays, it is possible to conceive a new type of GIS project based on a client-server approach using a relational spatial database where all data (e.g. geometric, alpha-numerical) can be collected together and directly managed by the client application on the WEB. Open Source solutions today offer more flexible instruments to manage this process compared with commercial solutions.

4.9 Technical performances

Up to now only economic and management comparisons have been described: a first analysis allows the practical parity between the two possible solutions to be understood. Many experiments conducted by the authors and or published in international literature show that this parity can also be appreciated as far as the offered technical performances are concerned.

The first test consists of comparisons between MapServer and ArcIMS (by ESRI) on the answering speed of server applications in publishing on the net spatial data. The used set of geometric data consists of 198.000 features with a number of vertexes ranging from 4 to 68.000 (average number of vertexes 80), 260 Mb of the shape file and 9.2 Mb of the dbf file. The system is asked to display different zoomed maps in a 550 x 400 pixel window. The application server is a Xenon DP2.8 GHz (1.5 Gb RAM) bi-processor and Linux OS: the data are stored in a Sun Solars server equipped with a Ultra Spare III 1 GHz (2 Gb RAM) bi-processor.

<table>
<thead>
<tr>
<th># area test</th>
<th># Polygons</th>
<th># Vertexes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>65855</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>66007</td>
</tr>
<tr>
<td>3</td>
<td>71</td>
<td>66387</td>
</tr>
<tr>
<td>4</td>
<td>562</td>
<td>79816</td>
</tr>
<tr>
<td>5</td>
<td>2645</td>
<td>138323</td>
</tr>
<tr>
<td>6</td>
<td>12715</td>
<td>567691</td>
</tr>
<tr>
<td>7</td>
<td>76864</td>
<td>5248736</td>
</tr>
<tr>
<td>8</td>
<td>194801</td>
<td>16331074</td>
</tr>
</tbody>
</table>

Table 1: Entities involved in different areas

<table>
<thead>
<tr>
<th>PSTGW</th>
<th>PSTGS</th>
<th>SDE</th>
<th>SDO</th>
<th>SDO-OGR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.099</td>
<td>0.890</td>
<td>1.060</td>
<td>0.804</td>
</tr>
<tr>
<td>2</td>
<td>5.186</td>
<td>0.739</td>
<td>0.998</td>
<td>0.799</td>
</tr>
<tr>
<td>3</td>
<td>5.221</td>
<td>0.824</td>
<td>1.055</td>
<td>0.741</td>
</tr>
<tr>
<td>4</td>
<td>5.247</td>
<td>1.165</td>
<td>1.133</td>
<td>0.830</td>
</tr>
<tr>
<td>5</td>
<td>5.601</td>
<td>2.081</td>
<td>1.418</td>
<td>1.210</td>
</tr>
<tr>
<td>6</td>
<td>8.065</td>
<td>12.758</td>
<td>3.568</td>
<td>3.609</td>
</tr>
<tr>
<td>7</td>
<td>32.827</td>
<td>31.852</td>
<td>24.620</td>
<td>26.290</td>
</tr>
<tr>
<td>8</td>
<td>86.357</td>
<td>100.627</td>
<td>76.158</td>
<td>79.207</td>
</tr>
</tbody>
</table>

Table 2: MapServer results (time expressed in seconds)

<table>
<thead>
<tr>
<th>SDE</th>
<th>SDO</th>
<th>SDO DC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.751</td>
<td>11.057</td>
</tr>
<tr>
<td>2</td>
<td>0.765</td>
<td>11.055</td>
</tr>
<tr>
<td>3</td>
<td>0.835</td>
<td>11.371</td>
</tr>
<tr>
<td>4</td>
<td>0.988</td>
<td>11.738</td>
</tr>
<tr>
<td>5</td>
<td>1.607</td>
<td>13.970</td>
</tr>
<tr>
<td>6</td>
<td>3.755</td>
<td>27.185</td>
</tr>
<tr>
<td>7</td>
<td>26.093</td>
<td>TIME-OUT</td>
</tr>
<tr>
<td>8</td>
<td>47.526</td>
<td>TIME-OUT</td>
</tr>
</tbody>
</table>

Table 3: ArcIMS results (time expressed in seconds)
commercial and Open Source solution. These results have been confirmed by other tests developed by Refraction Research Inc. (PostGIS and uDig Open Source programs developer). Figure 1 shows the MapServer and ArcIMS delay times when managing different numbers of entities obtained at the Ministry of Sustainable Resource Management (British Columbia, Canada).

![Figure 1: MapServer and ArcIMS delay times](image)

A second test was performed on geographic coordinate conversion tools. Open Source solutions pay a great deal of attention to the data reference systems. Mainly PROJ library is used above all to define the involved reference systems and the transformations from one system to the others; this possibility is also present available in commercial solutions. The test data are 90 raster map sheets provided in NTF Lambert II etendu by IGN (France) in a 1:25.000 scale (23 Mb each map sheet). These data were converted into a geographical coordinate system based on WGS84 Datum. A traditional 7-parameters spatial transformation algorithm was used. Operating on a standard PC (Pentium IV 2.8 GHz, 1Gb RAM) GDAL utilities operates the transformation in less than half an hour. On the same hardware ArcGIS is not able to perform the transformation. ArcGIS performed the test on a bi-processor but a couple of hours were necessary.

4.10 Final considerations

Summarizing the previous analysis it can be concluded that Open Source solutions are a valid alternative to commercial platforms both from an economic and a technical point of view. Besides, other social and economic advantages can be found for the Open Source solutions.

It is well known that Open Source does not mean free of charge! A complete development can be more expensive than a commercial solution, but more elements have to be taken into consideration in order to perform a correct economic analysis. Open Source solutions can be developed by using and promoting local expertise, therefore the costs can also be justified from a social point of view.

5. OPEN SOURCE AND STANDARDS

The connection between Open Source and standards is very close. The new APERI initiative, which involves Open Source projects and private Companies (e.g. IBM, Cisco, Sun, etc.) aims at defining an Open Source platform that is free from royalties for the management of business archives. The main goal of this important initiative is the creation of an open standard for the management of storage infrastructures. The final objective is to overcome API (Application Programming Interface) limitations.

One of the most important initiatives of the Open Source inside standardisation is the Open Gis Consortium (OGC) a no-profit international organisation which develops standards for spatial and location-based services. OGC, using the consensus-programs, collaborates with governments, private companies and universities in order to create programming interfaces which allow the development of interoperable applications.

Today OGC is made up of more than 250 members; the first private company involved in OGC was INTERGRAPH. The development of software applications based on OGC standards allows the user to benefit from a wider WEB solution environment.

WMS (Web Map Service) was the first standard to be developed by OGC and it provides the dynamic production of spatial referenced maps starting from spatial data. GML (Geography Markup Language) describes the spatial information code using XML format. GML allows the transportation and the storage of the modelled spatial data. WFS (Web Feature Service) defines an interface which allows access and manipulation of the spatial entities using the http protocol as a distributed platform. By using these interfaces, it is possible to create new features, to cancel and/or update spatial entities and to select or query features using spatial or not-spatial keys.

![Figure 2: Example of a geometry table for a polygon by using SQL and a textual representation of the entities](image)

Interoperability inside the spatial analysis requires a common declaration of the geometric features. The "OpenGIS Simple Feature Specification of SQL" defines an SQL standard in order to record, search, select and update simple spatial data by using Open Database Connector API. Simple spatial data are defined as 2D geometric components formed by segments and vertexes. The storage of these features is performed using geometric columns inside a generic Database table. Figure 2 shows a simple geometry record according to OGC specifications.

6. BASIC OPEN SOURCE TOOLS FOR SPATIAL DATA MANAGEMENT

A brief summary of the best tools for managing spatial data is provided in this section. These tools are the basic instruments necessary to build-up Spatial Information Systems under Open Source specifications. The first interesting tools are those of the PROJ.4 (http://proj.maptools.org) family: these are able to convert geographic into Cartesian coordinates and vice-versa. Originally written in Fortran language, they have been translated into C programming language.
GDAL (Geospatial Data Abstraction Library – http://www.gdal.org) is a set of tools which allows conversion between different spatial data formats (e.g. ENVI, ESRI, GEOTIFF, JPEG2000, etc.). The GDAL library is equipped with a set of utility software which is able to perform simple activities on raster files (e.g. copy, overview production, resample, contour line creation from a DEM, etc.).

OGR (http://ogrgis.maptools.org) allows different vector formats to be read (e.g. shp form ESRI, tab form MAPINFO, etc.). OGR tools are distributed with a set of basic routines to convert vector formats.

GD tools (http://www.boutell.com/gd/) allow a dynamic creation of raster images in the most common formats (e.g. PNG, JPEG, GIF, etc.).

ShapeLib (http://shapelib.maptools.org) is a simple API which allows the writing of simple software able to read and write ESRI ArView Shape files and the related database files (dbf).

JTS Topology Suite gives a model for spatial objects and basic functions of geometric analysis (http://www.vividsolutions.com/jts/jtshome.htm). It supports the spatial features of the OGC Simple Feature Specifications. Binary analysis (e.g. equal, disjoint, intersect, touch, cross, within, contain and overlap) are provided. JTS also provides the main spatial analysis methods.

Quantum GIS (http://qgis.org) is the most commonly used Open Source instrument for desktop visualisation of raster and vector data. Other similar projects are GVSIG (http://www.gvsig.gva.es) and THUBAN (http://thuban.intevation.org).

7. DESKTOP GIS OPEN SOURCE SOLUTIONS

The previously described tools and library can be used to build-up and integrate a desktop GIS project developed under Open Source specifications.

The most important and frequently used tools are briefly described in the following.

JUMP (http://www.jump-project.org) is an application which can manage and visualize spatial data. It includes several GIS tools and can be completely customised by implementing an ad-hoc plug-in developed in the JTS environment.

A large number of geometric editing tools are directly provided in Jump applications and spatial analysis. Spatial conversions and geometry validation are possible in a friendly and easy to learn environment.

UDig (http://udig.refractions.net) is an application in which the users can create, acquire, visualize and edit spatial data on a platform where the programmers can develop new applications.

GRASS (http://grass.itc.it) is the oldest Open Source project in desktop GIS solutions. This platform is now used in many European universities for teaching purposes. Its long history justifies the huge set of applications developed and distributed to all the users.

8. WEB GIS OPEN SOURCE SOLUTIONS

Open Source platforms provide all the instruments necessary to build-up a WEB solution for a Spatial Information System that can be completely used interactively.

8.1 Spatial application environments

One of the most diffused instruments for the development of spatial applications on the WEB is MapServer (http://mapserver.gis.cnm.edu) which managed by the University of Minnesota.

MapServer has been implemented using Shapelib, FreeType, Proj.4 and GDAL/OGR Open Source projects. It runs on any kind of operating system (e.g. Linux/Apache, Unix, Microsoft, Mac). MapServer integrates MapScript which allows the development of applications using the most common programming languages (e.g. C, Perl, Python, PHP, Java, C++). For example, using a Perl Database Interface Module, it is possible to integrate data coming from different DBMS (e.g. Oracle, Sybase, MySQL, PostgreSQL/PostGIS) with information from traditional GIS software.

Another interesting programme is GEOSERVER (http://geoserver.sourceforge.net) which provides an interoperable infrastructure that is useful to share spatial information according to Open Gis Consortium Standards. One of the main advantages of GEOSERVER is the possibility of upgrading, editing and inputing spatial data using the Transnational WFS protocol.

A third possible solution is Degreer (http://deegree.sourceforge.net) which provides instruments to build-up a Spatial Data Infrastructure by implementing Open Gis Consortium and ISO/TC211 standards. Degreer guarantees full integration with commercial and Open Source GIS platforms. Developed in Java environment it can be managed by different Operating Systems (Microsoft, Mac, Linux, Unix, etc.).

8.2 Relation databases with spatial support

This is the best kind of database instrument to manage and record spatial data and can be accessed both by desktop and WEB GIS solutions.

The two most interesting products are PostGIS and MySQL.

PostGIS adds the spatial support the well known PostgreSQL relational database (http://postgis.refractions.net). It has been equipped with simple desktop interfaces to allow an easy visualisation and editing of the recorded data. The Open Gis Consortium specifications are all adopted and advanced topological structures (e.g. coverage, network, surface, etc.) are supported.

The spatial data are managed using GIST (Generalized Search Trees) indexes in order to speed-up the search and the section of the data inside the database.

All spatial data can be converted in graphic form using the OGR tools (http://ogrgis.maptools.org).

MySQL (release 4.1 – http://dev.mysql.com) allows the generation, recording and analysis of spatial entities; the testing of spatial relations between geometric features has not yet been developed.
8.3 General considerations

The Cultural Heritage is generally managed by public administrations and/or organisations. A commercial approach to GIS for CH documentation means several licenses must be bought and maintained and only a reading of the collected data can be distributed through WEB.

The Open Source approach allows the free distribution of a self-built solution and the possibility of managing the system via WEB: in many cases specialists involved in a CH documentation project come from different regions or Countries and Open Source solutions allow a truly widespread operability without any extra-costs.

The present development also allows the conversion of old solutions, based both on commercial and Open Source desktop applications, into Web based Open Source applications.

9. CONCLUSIONS

The use of GIS technology in CH documentation is today mandatory.

The basic research can be considered closed and now it is necessary new efforts to be done to fix general rules for a correct implementation.

These rules have to respect the standards used in Land planning applications as much as possible, but it should be considered that each CH object requires an ad hoc solution.

It will be necessary for desktop solutions to be replaced by WEB solutions in order to allow the dissemination of the results and to increase the possibility of international collaboration.

Commercial and Open Source platforms can be profitably used thanks to the high degree of customisation they allow.

The Open Source approach offers more interesting advantages than commercial approach such as the growing of local expertises and the possibility of sharing the developed solutions without any costs.

After a long period of self made projects the RecorDIM initiative tried to formalise the GIS approach in Cultural Heritage documentation. The results achieved in this three year project will be used to produce guidelines on “Generic Templates for the management of heritage places” this project concerns a collaboration between the authors and Prof. Gutliz Bilgin of the University of Ankara (Turkey).

The RecorDIM initiative will find in CIPA activities a natural proceeding in order to fix and test a new approach in CH documentation which will be open to all Countries due to the limited economic costs.

REFERENCES


