Frank Henze – Heike Lehmann – Wolfgang Langer

CISAR – A Modular Database System as a Basis for Analysis and Documentation of Spatial Information

Abstract: For archaeological research work on large survey areas and complex building structures, there is a need for specific spatial information systems that offer comprehensive combination of geometric data with thematic information and additional components. Three years ago, the development of a web-based database and geo-informational system was started at the BTU Cottbus in cooperation with the German Archaeological Institute in Berlin for an interdisciplinary research project in Baalbek. Due to the complexity of the multidisciplinary research, a modular system structure was necessary to allow an integrative examination and analises of results from several particular projects. This modular design allows for adaptation to other research projects on building archaeology at the BTU Cottbus; in addition to the application and improvement of existing modules, new database modules could be developed and integrated into the system. The modular design of the database system and its basic modules will be introduced and the interaction of specialized modules with the basic modules and their multidisciplinary application will be examined. Furthermore, the development of specific tools for the creation of consistent graphical user interfaces for all modules and for standardized communication between database modules will be realized. The database system presented, CISAR, is the basis for project-specific solutions in the visualization and analysis of spatial relations within 2D and 3D GIS. The use of CISAR for a multi-scalar GIS as well as for a 3D building information system is presented in two additional papers.

Introduction

Efficient databases for storage, documentation, analysis and presentation of research results have become an important tool in the field of archaeology. The majority of the data exhibits a spatial reference and thus techniques of geo-informatics, which have been used successfully in other areas for many years, can also be applied for archaeological purposes (Heine / Henze / Riedel 2005; Heine 2006). CISAR is a modular information system for archaeology and research on building history which has been developed through cooperation between research projects at the Brandenburg University of Technology in Cottbus and the German Archaeological Institute (http://www.tu-cottbus.de/cisar).

The development of CISAR began with the need for a spatial information system for two different research projects: on the one hand, the archaeological investigations of the urban development of Baalbek in Lebanon (van Ess et al. 2003) and on the other, the research on building history of the Domus Severiana Palace on Palatine Hill in Rome (Hoffmann / Wulf 2004). Through the cooperation of the two projects, the primary project-specific solution for Baalbek was able to be developed and expanded towards the creation of a multi-purpose, modular, web-based information system.

Meanwhile, CISAR is also used for a project concerning the settlement history of the region of Triphylia in the Peloponnese, Greece. Several places are examined here with regard to kind, form, size and duration of settlement. For the documentation of urban structures, description of architecture and evaluation of archaeological finds, the existing modules of CISAR can be used with only a few adaptations. The time and expense needed to adjust the modules to the project-specific requirements is negligible in comparison with the expense that would result from conceiving a new database and / or a GIS solution.

Finally, CISAR will be used within the research project on the documentation of the building complex of the Pilgrim Cathedral in Santiago de Compostela in Galicia for the realization of a building information system.

Modular Structure of CISAR

The basis of the CISAR-Information-System is a modular database, consisting of basic and specific

modules for the storage and management of thematic information. The database is complemented by several tools for administration, documentation and error reporting. The system incorporates basic modules for the storage of photos, plans, literature references and other documents as well as specific modules such as databases for archaeological finds, for architectural fragments, for building construction units and a room inventory. Information and objects from the basic modules can be assigned to objects in the specific modules.

Extending the thematic database system with spatial information generates a web-based Geographic Information System (GIS) for classical 2D problems, or a 3D Building Information System for architectural research. Web-GIS methods will be applied for the historical urban research in Baalbek (Henze / Mönicke 2006). The visualization of the complex building structure of the Domus Severiana requires a 3D application developed on basis of the VRML standard (Heine / Brasse / Wulf 2006).

The development of CISAR adheres to the requirements of the several research projects involved. As a result, it has not been possible to plan the complete database design at the beginning of the implementation work. Rather, it had to be modified and extended several times during the development process. But this procedure has allowed for data input from the beginning of the research project and produced developer tools for rapid and simple adaptation and expansion of database modules.

Web-based Technology

Web-based databases and information systems work as a client-server architecture, and they offer some advantages over desktop databases. Traditional LAMP architecture, consisting of a Linux operating system, an Apache web server (http://www.apache.org/), a MySQL database (http://www.mysql.com/) and PHP as interface programming language (http://www.php.net/), does not require any client software other than an up-to-date Internet browser, and there is no need for a certain operating system on the user side. All database functions, like data input, editing or query, can be realized with a standard HTML interface. All users and all projects work with a consistent graphical user interface (GUI), and they all work on one consistent server-based dataset. For each research project using the consistent database modules, a customized database is available on the server. The access is password-protected for each project, whereby graduated access rights can be assigned within each project.

It is advisable to use the services of a professional data processing centre for all server administration and maintenance. Long-lived hardware and automated backups guarantee very high data security and a high server availability, now and in the future. Once installed, server hardware and software generally function for many years without any system changes, and so server-based applications and the associated data stay accessible over a long period. This performance for a productive system would hardly be achievable with a server administered from an isolated project workroom.

The client-server architecture presented can be established using only open source software. Standardised and well-documented technologies and formats for data and communication guarantee independence from any proprietary and commercial software. The established open-source software and operating system of the LAMP architecture guarantee high security and stability because bugs and errors can be corrected by the developer community within a short time. Moreover, the philosophy of most open source systems is security by design instead of security by obscurity. Last but not least, building up such open source client-server architecture saves license fees for commercial software.

Basic Modules

All projects work with identically structured basic modules which include photo, plan, measuring point and literature database. Only the project-specific list entries such as photographers, draughtsmen or toponymy have to be adapted.

Photo Module and Plan Module

Images and plans are described in differently formatted fields, mainly by entries from attribute lists, in order to enable queries of a maximum of content. Using both modules, photo or plan data can be uploaded to the server as image files. After a file is uploaded, a thumbnail picture for a list overview and a preview picture for the individual data sheets are produced on the server automatically. Data records from the photo and plan modules can be linked with objects from the specific modules. This allocation does not influence the content of the data record, and

so the same record can be linked with more than one object. Within the basic modules, documents are intended to be found using the search function; however, they can also be found through their allocation to objects of the specific modules.

Measuring Point Module

The measuring point database supplies information regarding coordinates, the kind of marking and the characteristics of the surroundings to enable the user to retrieve points easily. In addition, measuring sketches can be uploaded on the server and measuring points can be visualised and queried in a GIS.

Literature Module

The literature database allows the user to search for publications and their reviews by different aspects and keywords. Additionally, their location and lending status in up to five libraries can be stored.

Specific Modules

Archaeological objects are stored within specific modules, such as a building and room inventory or a database for architectural fragments.

Archaeological Module

One of the specific modules of CISAR is an archaeological database realized for the Baalbek project to describe finds from current investigations. The database has a strictly hierarchical structure, so that every find is connected with information about the excavation background. For recording finds, it is necessary to specify the find spot according to the following steps:

1. First an area inside the town is defined.
2. Inside the area, there are different find spots; one find spot usually is a part of a building.
3. One find spot is investigated by several soundings.
4. Inside every sounding, a stratum describes the smallest spatial unit from which the find came.

For the Baalbek research project, the finds themselves are separated into three groups: small finds, small clay finds and pottery. For every kind of find there is a specific datasheet for its description. For instance, there is a specific module for the description of the fabric of clay objects. Information about the find position and the type of find is stored within the code of the inventory number and of course drawings and photos can be linked with the objects.

Survey Module

A tool designed to describe all kind of finds was necessary for a recently begun regional survey in Baalbek as well as for the survey project in Triphylia. Unlike that of an archaeological excavation, the spatial background of a survey find cannot be exactly specified. Therefore, the archeological database was adapted to the needs of a survey project by breaking with the strict hierarchical structure. Within the survey module it is possible to assign "free finds" on different levels of the database. For the description of free finds, a universal datasheet is provided for a large variety of finds and the various aspects of their description. The core of the survey database is the previously described archaeological database, and all the levels of excavation backgrounds find their expression within the survey module:

1. The area of a survey defines a settlement, for instance the ancient town of Makistos in Triphylia.
2. A find spot specifies a part of the town, like the Acropolis of Makistos. From this level of specification on, finds such as architectural fragments, shards or stone objects can be allocated.
3. If the context is known in more detail, the find can be attached to a building, for example the Athena temple.
4. The find may even be associated with a part of the building, for instance the Peristasis of the Athena temple.

Room Inventory Module

Another powerful specific module of CISAR serves to describe building structures. The room inventory was primarily planned to document the imperial palace Domus Severiana in Rome, which has a complex structure over several floor levels (Brasse / Riedel 2006). With a little amplification, this module can be applied as a universal tool for capturing all building structures, because the database structure follows the logic of construction of a building. Every part of a building, including the foundation, walls, roof, rooms, horizontal elements or openings, can be described. With the addition of information about the location, cadastre and inhabitants of the building, the module is also applicable for research on

settlements. Furthermore, it can be combined with the archaeological module. The room inventory module is presently used on two very different projects: the description of a Roman imperial palace and the investigation of traditional Ottoman buildings in Baalbek.

1. The area describes the investigated part of town in Baalbek and the palace area on the Palatine Hill in Rome.
2. The next level presents a building complex – the Domus Severiana palace or the Domus Augustana currently analysed for the Rome project. In Baalbek, a building complex means a single house; here the description includes ownership, cadastral number, use, state of conservation and so on.
3. From this level of specification on, descriptions of building parts with their materials, construction techniques and equipment follow the same rules in both projects. Only the list entries have to be adapted to project-relevant content. Furthermore, the relationship between building parts, like rooms and walls or neighboring rooms, can be defined.

Query Functionality

The great advantage of the use of databases for data collection and storage is the opportunity for complex analysis. In CISAR, characteristics of and relationships between objects recorded in the databases can be searched with the help of a complex query module. This module enables the user to create dynamic queries within a web interface.

First, the user selects the appropriate specific module of the database for his special requests. Various demands can be combined through logical operations to formulate really complex queries. Single requests can be combined by "and/or" constructions and various operations such as "contains", "higher/lower than" or "is/is not equal to". The results of a query are shown in the form of a list, which can be downloaded and saved as CSV file (comma-delimited) for further processing.

For visualization and further analysis, individual methods are built for every project: for the Rome project, the results of queries on the room inventory can be visualized in a three-dimensional VRML model, while the Baalbek project works with a two-dimensional Web-GIS for analysis of data that comes from the various modules. Both solutions will be presented in short papers within this publication.

Other Modules and Tools

In addition to the database modules, a set of organising tools is available in CISAR. These tools have the same structure, but the content is specialized for each project:

- Users are able to configure their individual profile within the administration module.
- There is a tool for error handling and adaptation requests.
- A local wiki, based on wikipedia technology, serves as an online user manual.
- A password-protected weblog can be used for internal project discussions.

Document Management Module

Currently, an independent Java-based document management system (DCMS), developed at the department for building computer science at the BTU, is used for central document storage. Here, several documents are stored on the data server, assigned to different folders and provided with additional information such as author, type of document or keywords. The documents can be searched either by browsing the folder structure or by a query model (Henze / Lehmann / Fischer-Genz 2005). Current plans are to implement an integrated document management module for central data storage within the CISAR system, so that documents can be linked with objects in the specific modules just as they can with data of the basic modules.

Implementation

For the implementation of the database system, several tools are used or were created for efficient programming.

ippfp Interface Library

From the beginning of the database implementation, the PHP library "ippfp" (http://www.ippfp.org/) was used to guarantee a consistent graphical user interface for all the different modules. The programmers are using the ippfp application programming interface (API) to simply create all the HTML interactions of a module. This means a manual HTML implementation for each single module interface is unnecessary.

Module Generator

At the beginning of implementation, new modules had to be created, whereas many code parts, functions and routines recur in every module. So the process of module generation can be mostly automated with only a small amount of input about the data types and connections used. A simple XML template file containing all the necessary information about a module is the input for the Module Generator. The result is a raw version of the new module, consisting of the PHP code and the attached SQL-table(s).

Module Interface

Functions for insertion, editing, viewing and deletion of base objects within a specific module are similar for all modules. Therefore, consistent methods for module communication were developed in the form of an application programming interface. This simplifies the integration and use of basic modules within either existent or newly developed specific modules.

Source Code Version Management

At present, several students and engineers are working on the development of CISAR. To avoid conflicts among different source-code versions, a version management system using SVN subversion (http://subversion.tigris.org) was installed on a special source code server. When starting work, the programmer loads the current code version from the SVN server to his local computer and works on a local copy of the source code. After finishing the implementation, he synchronises his work with the server version of the source code, whereby only an incremental code storage occurs, and every previous code version can be reconstructed and all changes can be tracked.

Conclusions and Prospects

The web-based information system CISAR, which has been utilised successfully for very different projects in archaeological fieldwork and building archaeology, has been developed within the last two years through the close cooperation of archaeologists, building researchers, geodesists and computer scientists.

The modular structure of CISAR makes it possible to simplify, adjust and extend the system with new specialized modules. The conceptual design and development of the system are currently oriented to the requirements of three research projects in which the BTU and the German Archaeological Institute are involved. Through the integration of further projects, the system can be advanced to become a tool applicable to very different requirements in historical research.

General standards for the documentation of cultural heritage objects (like recorDIM (http://extranet.getty.edu/gci/recordim/) or CIDOC CRM (http://cidoc.ics.forth.gr/)) have to be taken into account, as well as the information exchange with other established specialised information systems (like the Arachne database in Cologne).

Until now, the implementation of CISAR has been realised by scientific assistants and students of the BTU, whereby the incorporated research projects bear the costs. It has to be reconsidered whether the CISAR project can be continued as an open source project if the system is used by additional research projects. The main decision criteria is the problem of which institution can coordinate further development and whether a community of active developers exists that can cooperate in its advancement.

References

Brasse / Riedel 2006

C. Brasse / A. Riedel, Ein Bauwerksinformationssystem für die ‚Domus Severiana' und das Gartenstadion auf dem Palatin – Möglichkeiten und Grenzen. In: A. Riedel / K. Heine / F. Henze (eds.), Von Handaufmaß bis High Tech II – Informationssysteme in der historischen Bauforschung (Mainz 2006) 93–99.

van Ess et al. 2003

M. van Ess (with contributions from T. Bunk / V. Daiber, B. Fischer-Genz / F. Henze / K. Hitzl / F. Hoebel / B. Ritter / H. Wienholz), Archaeological Research in Baalbek. A preliminary report on the 2001-2003 seasons. Bulletin d'Archéologie et d'Architecture Libanaise (BAAL) 7, 109–144.

Heine / Henze / Riedel 2005

K. Heine / F. Henze / A. Riedel, Raumbezogene Informationssysteme für Bauforschung und Archäologie. Forum der Forschung – Wissenschaftsmagazin der Brandenburgischen Technischen Universität Cottbus 18, 2005, 91–98.

Heine 2006

K. Heine, Möglichkeiten und Grenzen von Geoinformationssystemen in der historischen Bauforschung. In: A. Riedel / K. Heine / F. Henze (eds.), Von Handaufmaß bis High Tech II – Informationssysteme in der historischen Bauforschung (Mainz 2006) 93–99.

Heine / Brasse / Wulf 2006

K. Heine / C. Brasse / U. Wulf, WWW-Based Building Information System for "Domus Severiana" Palace at Palatine in Rome by Open Source Software. In: M. Ioannides / D. Arnold / F. Niccolucci / K. Mania, VAST 2006 – Joint Event of VAST / CIPA / EG / EuroMed (Aire-la-Ville 2006) 75–82.

Henze / Mönicke 2006

F. Henze / T. Mönicke, Ein modulares OpenSource GIS für archäologische und baugeschichtliche Forschungen. In: C. Clemen (ed.), Entwicklerforum Geoinformationstechnik 2006 – Junge Wissenschaftler forschen (Herzogenrath 2006) 67–76.

Henze / Lehmann / Fischer-Genz 2005

F. Henze / H. Lehmann / B. Fischer-Genz, Development of an Internet-Based Information System for Archaeological Research and Studies on Urban History in Baalbek / Lebanon. 20th CIPA-Symposium (Turin 2005) 849–854.

Hoffmann / Wulf 2004

A. Hoffmann / U. Wulf, Die Kaiserpaläste auf dem Palatin in Rom: das Zentrum der römischen Welt und seine Bauten (Mainz 2004).

Frank Henze
Wolfgang Langer

Brandenburg University of Technology (BTU)
Chair of Surveying
Postfach 101344
03013 Cottbus, Germany
henze@tu-cottbus.de
awolfml@web.de

Heike Lehmann

Brandenburg University of Technology (BTU)
Chair of Building Archaeology
Postfach 101344
03013 Cottbus, Germany
Heike.Lehmann@TU-Cottbus.de