Developing an Information System for Archaeological Sites and Monuments - Data Model and Construction

David Haskiya
National Heritage Board
Box 5405
114 84 Stockholm, Sweden
Phone: +46 8 5191 8511 - Fax: +46 8 667 89 43 - E-mail: david.haskiya@raa.se

Abstract: The Swedish National Heritage Board is currently developing an information system for archaeological sites and monuments. The system is primarily intended for cultural resource management, forest protection management, physical planning and academic research.

The project is now in the second stage, which is composed of construction and testing of the information system and the drafting of a model for the administration and maintenance of the system.

Details are given on the modelling, design and construction of the system. Further topics are the classification of sites and monuments in the system and the use of XML for queries and data transfers.

Key words: information system, GIS, archaeological sites and monuments, data modelling, XML, meta-data

Project background

The topic of this paper is a development project where the aim is an information system for archaeological sites and monuments. The paper will elaborate on the modelling and design of the information system. It also includes a short background to the project and a brief description of Swedish cultural heritage management. The development project is carried out at the National Heritage Board, which is the central authority in Sweden for matters concerning the cultural environment and cultural heritage.

Existing information and databases

Systematic surveys and the Register of Ancient Sites and Monuments

In 1938 a nationwide survey for archaeological sites and monuments was initiated by the National Heritage Board (NHB). The reason for this survey was a government decision on mapping the landuse of the country. Archaeological sites and monuments protected by law were to be included on the Land Use Map published by the National Land Survey (NLS).

Information on the sites, in the form of geographic positions on maps and descriptions in notebooks was gathered in the Register of Archaeological Sites and Monuments (in Swedish "Fornminnesregistret"). Information from subsequent surveys, archaeological investigations and reports on finds made by private citizens, have since been added to the Register. The Register is still maintained by the National Heritage Board.

Today the Register contains information on roughly 400 000 locales each comprising one or more sites or monuments. Not all of the sites and monuments in the Register are protected by the Cultural Monuments Act; the Register also includes historically/ archaeologically interesting sites which are not protected by law.

Databases at the NHB and NLS

In the middle of the 1980's the National Heritage Board began digitizing the descriptions in the Register. The geographical information has been, and continues to be, partly digitized by the National Land Survey (NLS).

Attribute data was stored in an Informix database, "Forn", at the NHB, but the geographical data was stored at the NLS primary databank ("Geodatabanken - alfa") for Sweden. The fact that two different applications were used to register data in two different databases has made integrated computeraided analysis and management of the data complicated and time-consuming.

In 1998, as the past efforts of the National Heritage Board were never brought to a satisfactory conclusion, the NHB was given a commission by the government to solve the problem with digital access to this kind of information. In 1999 a development project with the aim to construct a new information system was initiated. The project will end in 2002 when the information system will be gradually deployed at the National Heritage Board and at the different County-council Boards.

Modelling and design of the information system

Requirements analysis

Modelling and design of the system really began when the project team travelled around the country interviewing the potential end-users about their wishes and demands for an information system on archaeological sites and monuments. Questionnaires were also sent to all stakeholders and end-users in order to get their input.

The requirement analysis then proceeded by drafting a "requirements document" through a number of brainstorming sessions in a workgroup consisting of members of the development project and representatives from different actors in Swedish cultural heritage management.

The resulting "requirements document" deals with all aspects of the system e.g. data content, possible queries, map background and system response times. This document is at the core of all modelling and design of the system. It will also be used as the basis for the acceptance testing protocol.

The specialized data model

Modelling continued with conceptual- and data modelling – the two were initially done in parallel, but the two workgroups merged after a few modelling sessions. The resulting data model consists mainly of an object representing the archaeological site or monument. This central object has relations to the secondary objects, events and references. Much effort was put into the modelling of a classification system for the central object (see below).

The general data model

Parallel to the development of the information system for archaeological sites and monuments the NHB also initiated a project to develop an overall information technology strategy. As part of this effort different existing data models (e.g. models for archaeological excavations or historic buildings) inspired the development of a common data model for the NHB. As a result the data model for sites and monuments can now be considered as a specialization of the common data model.

Classification of sites and monuments

The sites and monuments recorded in the database will be classified according to a system devised as part of the development project. The classification-system is basically three-tiered - types of objects can belong to one or more category and features can be added to every type of object (see below for an example).

Category: "Graves"
    Type: "Cairn"
        Feature type: "Form"
            Feature value: "Long cairn"

The (meta-) database currently allows for 12 categories and 160 types of sites or monuments, but through the use of features (types and values) the 160 types can be sub-classified in a number of ways. Categories, types, feature types and feature values can easily be added or modified with the option of letting such changes affect the classification of all database-entries.

XML for queries, data and meta-data

The requirements analysis emphasized the need for a platform independent manner of data transfer between system and end-user (and also from system to system) – the system developers at the NHB quickly turned to the use of XML to satisfy these design demands.

The resulting XML-format is tailor-made to contain data and meta-data based on the NHBs common data model, but it can also be used to query the system as well as create entries in the database. The format also comprises tags referring to binary data sent in the same "package" as the XML-document.

Using GML (Geography Mark-up Language) for geographical data was considered, but the NHB decided to wait until this format is formally standardized and versioned. In the meantime geographical data will be sent as ESRI shapefiles.

Data flow

The system will exchange data mainly with the NLS, but will also receive data from the National Maritime Museum (NMM). The data received from NMM derives from their database on underwater remains. The data will be imported through a database-link (both the sending and receiving systems use the Oracle DBMS). The imported data is converted to comply with the data-model, allowing it to be sent from the data layer to application layer in the same XML-format as the other data.

The information system will automatically send the geographical data and some attribute data for all sites or monuments protected by the Cultural Monuments Act, as well the same data for a number of other types of sites or monuments traditionally printed on official maps, to the NLS (for inclusion in their Cadastral Map). The geographical data will be converted to the NLS´s proprietary data-format AutoKA-FF before transfer, but some complementary data will be sent as XML.

Landscape and cadastral data for the system will be stored on a remote server (served by the NLS) dynamically producing GIFs for the system's web-based client.

Users of the system will be able to interactively download data from the system in a number of different formats. Initially the possible formats are ESRI shapefiles, MIF/ TAB-files, DXF-files and XML-documents. These four optional formats cover the needs of the vast majority of possible users.

Finally, third-party applications can be "certified" to communicate with the NHB-system provided that they utilize the NHBs XML-format for queries and data transfers. In this way third-party applications can dynamically query the system e.g. an electronic request from a forestry company to the National Forestry Board for permission to harvest an area of forest

triggers a dynamic query from the National Forestry Board's system to the NHB's system. The NHB's system then returns the data for any protected sites or monuments within the area to be harvested, which the National Forestry Boards system in turn re-packages and sends on to the forestry company.

System architecture and construction

The layers: Databases, Data layer, Application layer and Clients

As part of the NHBs information technology strategy the decision was made to implement a system platform that would be adaptable enough to handle many different types of data, uses and users. The general system architecture is as follows:

Databases containing data, geographical data and user information is stored in an Oracle DBMS. The NHB uses Oracle Spatial and Intermedia to store geographical data and images.

The data layer handles incoming XML commands and data from the application layer, converting them from XML to SQL. Database responses are in turn converted from SQL to XML and sent on to the application layer.

The application layer will handle incoming data and commands sending them on to the data layer. Responses from the data layer are converted from XML to HTML if the client is the thin web-client or just passed on as pure XML if the request comes from the client for data entry. By accessing the information stored in a user session the application layer can utilize different XSLs to produce HTML layouts in the thin web-client. In theory the application layer could also use other XSLs to produce other sorts of output e.g. PDFs or WAP-files.

In this way the system can tailor-make different outputs depending on what type of user is sending the request. A pupil in an elementary school and a professional in cultural heritage monument can thus have different layouts on their browsers with different query capabilities, receiving different system responses, and yet communicate with the same database.

Two clients for the system will be developed. The first will be a thin (web-based) client for searching and extracting data, the second a "fat" client for data entry.

The thin client will not demand any plug-ins and will run on Internet Explorer 4.01+ or Netscape 4.5+ and will allow the user to query the system either through geographical queries, attribute queries or a combination of both. Depending on the user's rights, data can be interactively downloaded from the system in a number of different formats (see above).

Summary

In response to an increasing demand for information on archaeological sites and monuments in digital form the Swedish National Heritage Board has initiated a project aiming for the construction of an information system holding data of this kind. The system will store c. 1,000,000 data-entries and will be deployed from 2002 onwards.

The aim of the project has been to construct a system that is expandable and adaptable. In order to do this we have developed a conceptual model for archaeological sites and monuments that does not stop the results of research, or changing heritage management principles, from affecting and changing the model.

Further, the National Heritage Board has decided to construct a general system architecture that can function as a platform for many different applications, dealing with different sorts of data. The architecture designed emphasizes the separation of storing the data from presenting the data (in the clients) by using XML in communications between application and database.
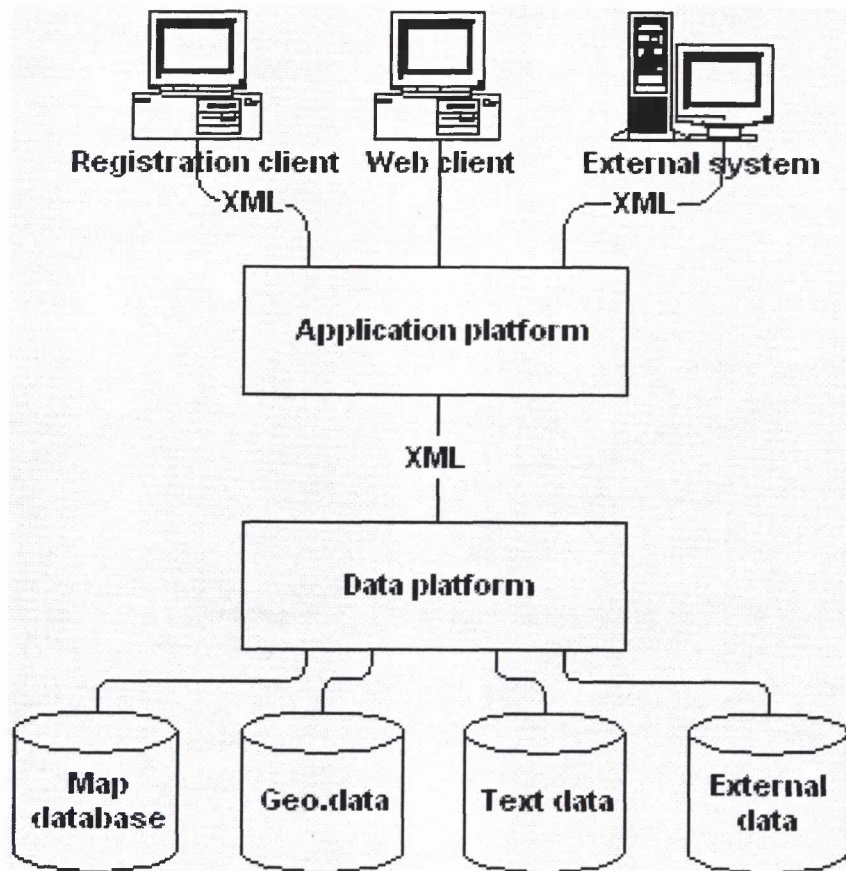
Figures

Figure 1. The architecture of the information system.