

# And XML-based open source solution for online/offline archaeological data storage and processing

PETKO STAYNOV, NIKOLAY GENOV

New Bulgarian University, Sofia, Bulgaria, petko@staynov.com

## ABSTRACT

*Arcada is a multilingual database management system destined to the storage and processing of archaeological data related to different types of artifacts: stone assemblages, pottery, etc. The analytical approach which stands at the bases of the system is the attribute analysis. The system is based on an open source technology and it is proposed in both an online and an offline version. The most original part of the system is the principle of its offline functioning which, thanks to XML format, allows the use of open source client/server technologies without the necessity to install server software on the client's machine.*

## INTRODUCTION

The presented open source solution is a database management system developed by the departments of Informatics and Archaeology of the NBU under the name of "ARCADA". The existing implementation of the system concerns different lytic assemblages from prehistoric sites on the territory of Bulgaria, Greece and Turkey. They have been made for the needs of different projects involving the participation of archaeologists from the NBU, the University of Skopje and the Institute for Mediterranean Studies, FORTH, Rethymnon, Greece. The goal of this paper is to present the existing versions of the system and to explain the ongoing developments and improvements.

## 1. MAIN PURPOSES OF ARCADA

The system has been developed with respect to the following goals:

- to allow the storage of data concerning individual archaeological artefacts permitting further quantitative processing (with SPSS, or other software products for statistical analyses);
- to allow the storage of generalised information about archaeological artefacts
- to avoid mistakes, ambiguities and misinterpretation of data by the imposition of data representation standards;
- to allow the exchange of data provided by different research teams;
- to allow the common processing of data coming from different archaeological sites;
- to allow the spatial representation (on a digital map) of gathered and processed data.

## 2. TARGETED ARTIFACTS

Normally ARCADA can be used with any kind of archaeological artifacts. The modules developed so far concern assemblages of stone artifacts, but the open architecture of the system allows the users having the necessary rights to set it for any other kind of material. The only limitation could result from analytical approach which stands at the bases of the system – the attribute analysis.

## 3. BASIC CHARACTERISTICS OF ARCADA

- client/server architecture – allows the sharing of the system functionalities by a large number of users, as well as the storage of larger amounts of information coming from different sources;
- multilingual interface – as all the text information (data and interface elements) is encoded, it is easy to add new languages both to the interface and to the core data; the language specific information is stored in XML schemes;
- different levels of access – the system allows an unlimited number of user types with different access permissions; the user types and the corresponding rights are set by the system administrator;
- multimedia features – the system allows the integration of static and dynamic images, and audio files; a specific module, developed for a Greek-Bulgarian project, allows the visualization of data on digital maps;

- support of numerous input and output data formats – xls, csv, etc.
- offline/online mode of functioning.

#### **4. BASIC FUNCTIONALITIES OF ARCADA**

- entering of data through a friendly interface, adaptable to the needs of the different types of users
- preparation of data for quantitative processing
- online publication of data
- online aid for translation of specialised documents.

#### **5. MAIN TYPES OF POTENTIAL USERS**

The system was especially designed for the following types of users:

- field workers – they can use the system mainly for the documentation of the results of their field research, but also for comparing data and preparing findings for publication;
- archaeologists working in laboratory conditions – they can also use the system as a documentation tool, and as an interface tool for further quantitative analyses;
- scholars doing interdisciplinary research – the system allows the preparation of data for different kinds of further processing;
- university lecturers and students – the possibility to associate pictures to the artefacts and to present generalised data on digital maps makes ARCADA a suitable tool for teaching/learning archaeology;
- museum workers – the system provides most of the functionalities which are usually present in a museum information system;
- virtual exhibitions curators – the XML description of data allows a relatively easy integration of the artefacts to different kind of virtual exhibitions.

#### **6. THE BASIC MODULES OF ARCADA**

The system comprises two main types of modules:

- general modules – dedicated to general archaeological data – site description, absolute and relative dating, institutional context, copyrights, etc.;
- specific modules – concerning the different artifact assemblages and collections and the different descriptive and analytical approaches.

All the modules are open in manner – the users of the system can either add, or modify the already entered attributes and characteristics, depending on the rights they have been assigned with. However, the CIDOC compliancy which has to be implemented at a further stage will impose a certain number of constraints.

#### **7. GENERAL SITE INFORMATION**

The general site information module has been designed with respect to the Bulgarian standards of archaeological documentation. Its structure is open and can be adapted to any particular national regulation.

#### **8. GENERAL ARTEFACT INFORMATION**

In figure 2 are presented the general categories that are usually associated with an archaeological artefact. The scheme is not limitative and the list of the implemented categories can be appended by new ones. The only requirement to respect is to always associate the new categories or attributes to the existing ones using property links.

#### **9. SPECIFIC ARTEFACT INFORMATION – ARTEFACT DESCRIPTION**

The scheme bellow presents a part of the categories used in the attribute analyses of lytic artefacts. They take into consideration the type of raw material, the specific production technology, the artefact's physical aspects, the traces of use, etc. These categories (or attributes) are presented to the user under the form of controlled lists. In order to add new categories, or to edit the existing ones, the user has to have the appropriate rights. The new specific categories are derived or connected to the existing ones by means of a certain number of generation rules. One of the purposes of those rules is to ensure the system uniformity and to avoid categorical ambiguity – they make almost impossible the introduction of a new category which is in fact just a synonym of an existing one.

## 10. CIDOC CRM COMPLIANCE

One of the main aims of the CIDOC CRM is to be a guide to conceptual modelling in the domain of cultural heritage, but it is mainly related to museums. The ARCADA system is CIDOC CRM compliant inasmuch as:

- the categories in its general modules – site information, dating, etc. – follow the CIDOC CRM class hierarchy;
- the artefact properties in the specific modules correspond to the CIDOC CRM property hierarchy, or can be considered as extensions of this hierarchy;
- all new categories that can be created by the user of the system should be derived from the existing ones by generalisation/specification or by a property link.

## 11. ONLINE FUNCTIONING OF THE SYSTEM

The chosen technology is based on a MySQL database accessed through the Internet thanks to PHP scripts. The same existing implementations of the system are based on single server architecture but if needed it can be realized as a distributed application, too.

The system allows the definition of different roles depending on the access and action permissions.

### Typical roles:

Super-administrator – he is the main administrator of the system; he defines and assigns roles (especially with administrator's rights), creates modules and tables. He has not the possibility to edit and delete data entered by the other users of the system.

Administrator – he defines roles (operators of different levels), creates modules and categories, defines links and enters/edits/deletes categories, links and data.

Operator – enters/edits/deletes data according his rights.

## 12. OFFLINE FUNCTIONING OF THE SYSTEM

The offline version of the system is under development. It is dedicated to fieldworkers who have no online access. The use of an XML based technology will allow them to work with the same interface as if they were online, without having to install the server side software used in the online version – MySQL database, PHP, web server, etc. At the same they will be able to make all the modifications allowed by their respective role and to update the online system with them once they have been connected to the Internet. The following scheme presents the basic principles of the adopted solution.

## 13. CONCLUDING REMARKS

The ARCADA system is a powerful versatile tool for archaeological documentation. But its further improvement will depend mainly on its implementation into different archaeological sub-domains and its use with different analytical approaches. The authors' intentions are to test it with different subjects and languages, as well as with different software and hardware environments.

FIGURES

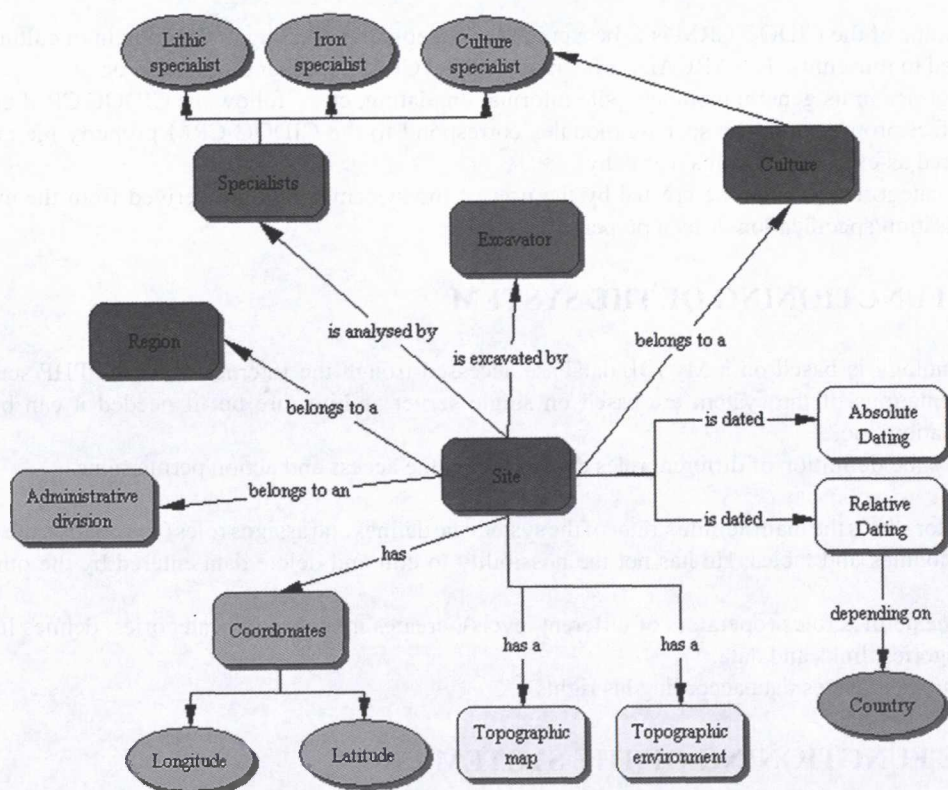


Fig. 1 – Main parameters of the general site information.

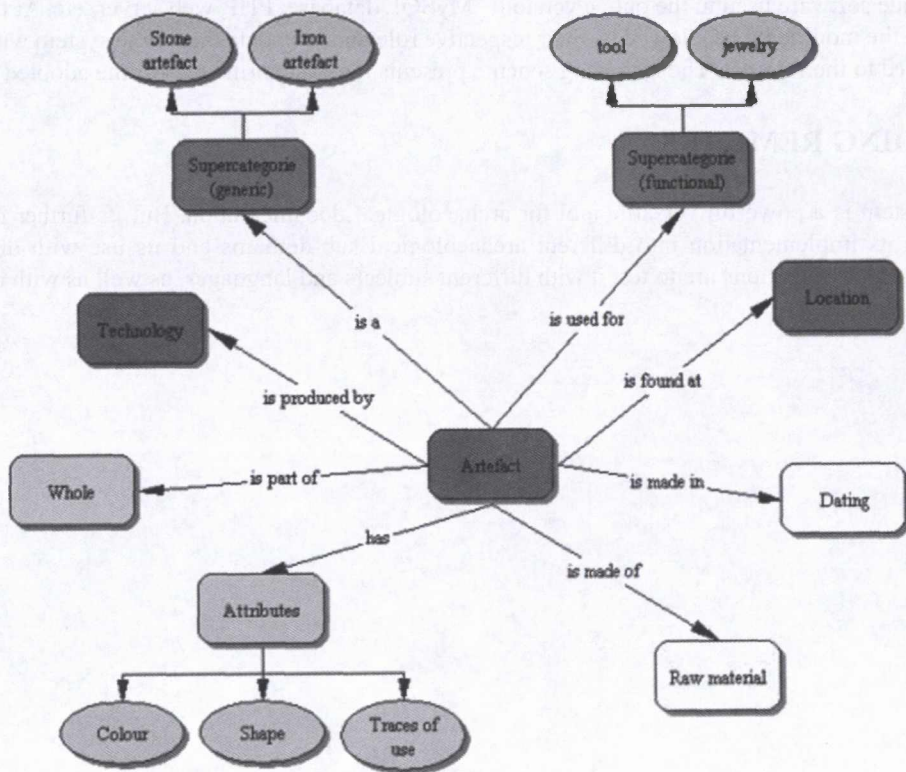


Fig. 2 – A conceptual scheme of the general artifact information.

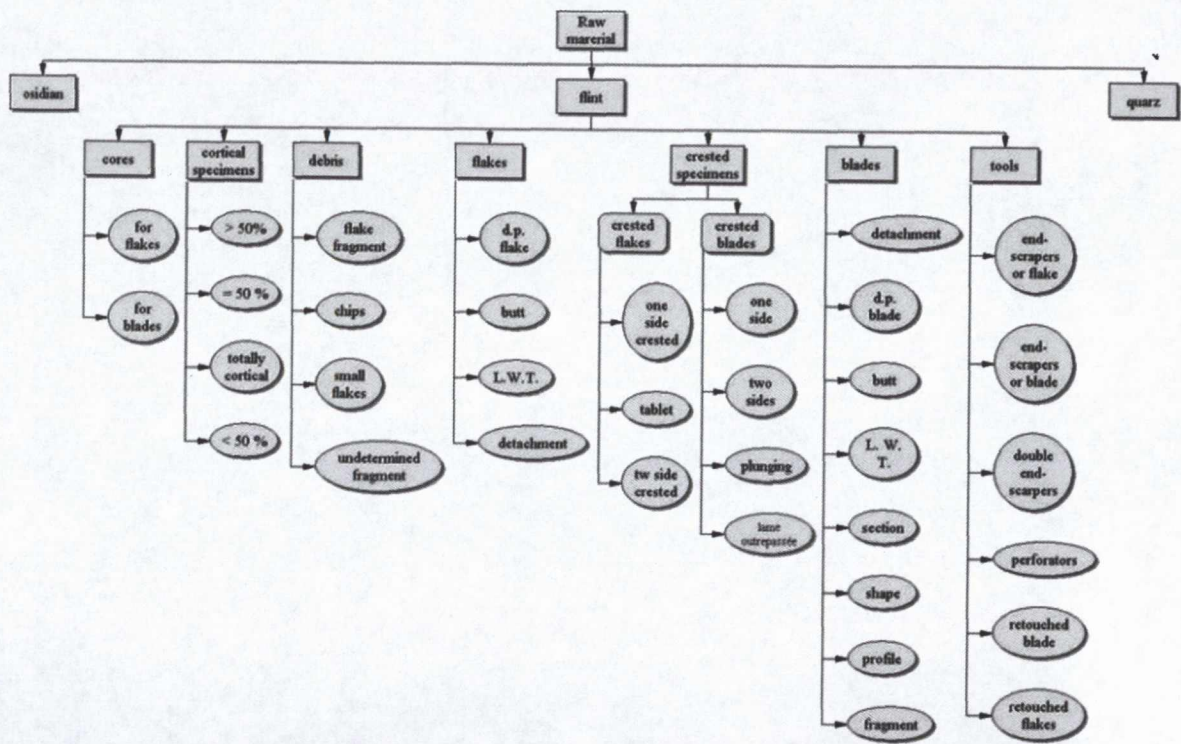


Fig. 3 – A tree of certain lytic artifact attributes.

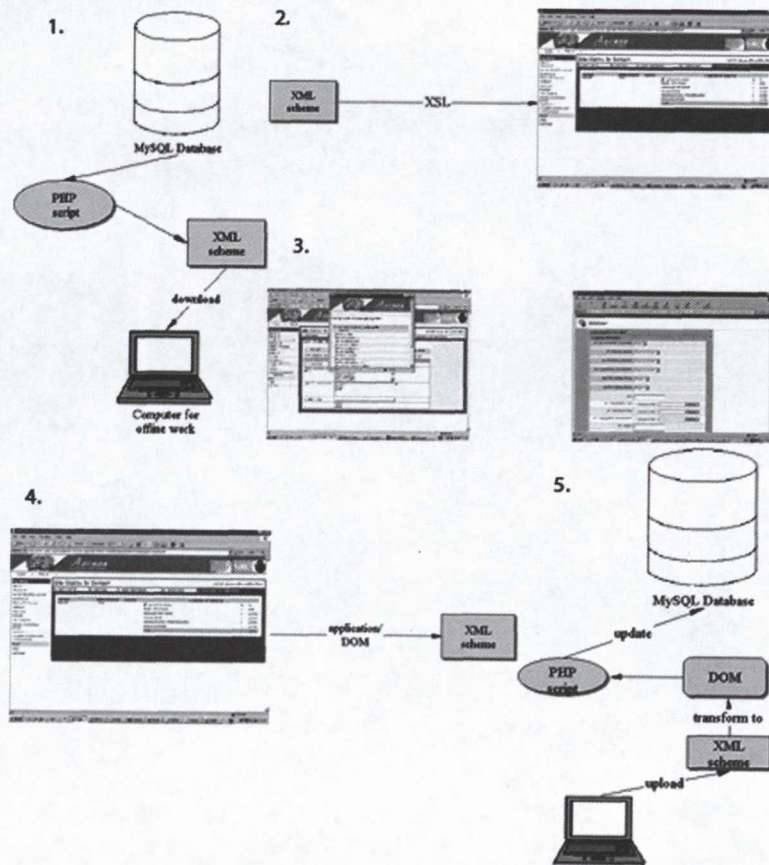


Fig. 4 – Basic steps of the offline functioning: 1. Download of a data scheme from the server or creation of a new scheme by means of a special application (HTML format – browser). The scheme is stored on the PC in XML format; 2. Visualization of the scheme in HTML format (XML – XSL – HTML); 3. Modification of the scheme (if needed) by the creation of new tables, fields or the definition of new values (HTML format – browser). Entering/editing/deleting data (HTML format – browser); 4. Saving the modifications (HTML interface – application/DOM – XML scheme); 5. Uploading of the scheme filled with data on the server (XML – DOM – PHP script – MySQL DB).