Digitizing Early Farming Cultures

Customizing the *Arches* Heritage Inventory & Management System

Edeltraud Aspöck/Anja Masur

OREA Institute of Oriental and European Archaeology
Austrian Academy of Sciences
Vienna, Austria
Edeltraud.Aspöck@oeaw.ac.at; Anja.Masur@oeaw.ac.at

***Abstract*—The aims of the project 'Digitizing Early Farming Cultures' (DEFC) are to create standardized and integrated research data of Neolithic and Chalcolithic sites and finds of Greece and Anatolia (c. 7000–3000 BC according to Greek terminology), two neighbouring and archaeologically closely related regions usually studied in isolation of each other. The data will be made available open access online ensuring compliance to standards in data production for data sharing (metadata and mappings) and interoperability with related initiatives. In this paper we will present first results of the project. Based on focus group meetings and analysis of our resources we created a conceptual data model for a site database. We will implement the data structure into the *Arches* Heritage Inventory & Management System and extend the *Arches* CIDOC CRM graph for the mapping of our data.**

***Index Terms*—Archaeology, site database, Neolithic Greece and Anatolia, *Arches*, CIDOC CRM**

I. Introduction

Greece and Anatolia in the Neolithic and Chalcolithic periods (c. 7000–3000 BC according to Greek terminology) are two neighbouring and archaeologically closely related regions which are usually studied in isolation of each other. As a result, research across the area currently suffers from fragmented data organized according to differing knowledge schemes (terminologies, chronologies), hindering collaborative research.

The project 'Digitizing early Farming Cultures' (DEFC) will provide standardized research data by integrating non-digital and digital sources from a series of research projects of the new research group Anatolian Aegean Prehistoric Phenomena (AAPP) at the OREA Institute (Austrian Academy of Sciences) [1]. AAPP studies sites and finds across the region, focusing on questions regarding technological and social changes, settlement patterns, exchange and sourcing of raw materials across the region. The data will be made available open access online ensuring compliance to standards in data production for data sharing (metadata and mappings) and interoperability.

In this article we will present first results and our current work in progress. Analysis of resources and focus group meetings led to specification of user needs and the creation of a conceptual data model for a site database. We will implement the model into a graph database incorporating international standards in cultural heritage management. We are now in the process of extending the CIDOC CRM graph that is part of *Arches* Heritage Inventory & Management System (HIP) for the mapping of our project data.

II. AAPP resources

A. Digital resources

Several site- and finds databases, including e.g. a MySQL site database on Western Anatolia (C. Schwall, ERC project 'Prehistoric Anatolia') that contains site data (GPS coordinates, dating, information on finds and archaeological features) and finds databases from Greece and Crete (E. Alram).

B. Analogue resources

'Ägäische Frühzeit' is a comprehensive publication with reports and analyses on excavations of Neolithic to the early Bronze Age (6500-2000 BC) sites in Greece. [3]. The Ceramics collection Fritz Schachermeyr is a 'teaching collection' of Mediterranean prehistoric ceramics with diagnostic Neolithic finds of the region (Fig. 1) [4].

Fig. 1. Image of 3D scan pottery sherd Schachermeyr collection (S. Štuhec).

III. Data model

In focus group meetings we discussed requirements for the system. Firstly, researchers wanted a site database that provides more detail than existing online open access site databases for the region [5]. It should be possible to query for specific information and to add information on 'closed contexts' (ensemble of finds deposited together at the same time, e.g. an undisturbed grave) to enable chronological comparison of sites. Another requirement was the presentation of typical – 'diagnostic' finds and archaeological features for a period/region/site – e.g. potsherds from the Schachermeyr pottery collection (Fig. 1).

978-1-5090-0048-7/15/$31.00 ©2015 IEEE

A conceptual data model was developed based on differences and similarities in content and structure of the existing databases and analogue resources (Fig. 2). No chronology for the whole region could be established but regional chronologies have to be used: Anatolia; southern and central Greek mainland including Cyclades; Thessaly; Macedonia/Thrace; Crete. Word lists with English and German terminology were created and we are in the process of creating thesauri and relating them to existing thesauri [6].

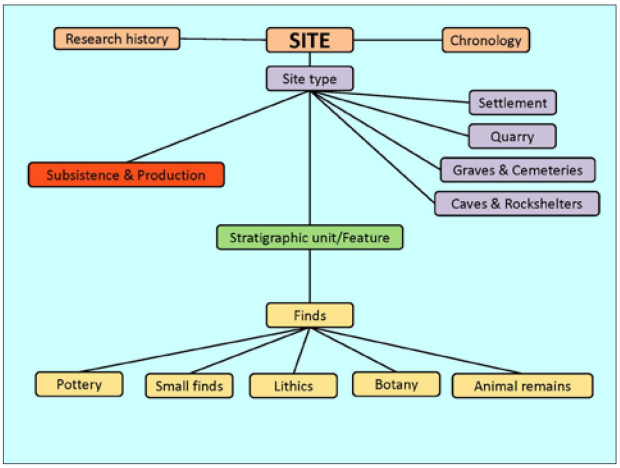

Fig. 2. Abstract of DEFC data model.

IV. Extending *ARCHES* CIDOC CRM structure

Arches is an open source software system designed for the heritage sector to inventory and manage all types of immovable cultural heritage. It is a graph database with the ontology CIDOC CRM integrated – the most widely used ontology for cultural heritage [7]. The *Arches* application 'Heritage Inventory Package' (HIP) can be customized to user needs and also includes a tool for creating thesauri, mapping features and a timeline. This allows to publish data online according to internationally accepted standards.

We are currently in the process of customizing the *Arches* CIDOC CRM structure for our data:

- Specification of the *Arches* default mapping to CIDOC: *Arches* is built to map as many data as possible and it is therefore often very general. A superclass may be used instead of a subclass, e.g. any archaeological site is mapped to 'Heritage Resource. E18'. In the mapping of our data we will be more specific to lose as little semantic expressiveness as possible.
- Add extensions relevant for archaeological data [8]: The *Arches* vocabulary is limited to a part of the core CIDOC CRM and we will need an extension to enable uploading of all our data in a CIDOC compatible way. For example, to represent our archaeological excavation data we will need CRMarchaeo.

V. Conclusions

The creation of a site research database of multiple archaeological regions with different research traditions, terminologies, chronologies that allows querying on a deeper level, e.g. for cross regional analysis poses the challenge of creating a data structure that fits all. Using a graph database that incorporates international standards will allow to organise our data closely to our model and it makes data interoperable for open access online publication. Customizing *Arches* for our site data will create a powerful tool which we will share with the wider community.

Acknowledgments

This project is supported by the Austrian Academy of Sciences initiative go!digital (ACDH 2014/22) and by the ARIADNE EU-funded project (FP7-313193). We want to thank Eva Alram and Christoph Schwall (AAPP research group) for their contributions to the data model and thesauri and Matej Durco and Hannes Pirker (ÖAW ACDH) for feedback and technical support.

References

[1] AAPP: http://www.orea.oeaw.ac.at/index.php?id=172 B. Horejs, "Proto-Urbanisation without Urban Centres? A Model of Transformation for the Izmir Region in the 4th Millennium BC", / E. Alram-Stern, "Times of Change: Greece and the Aegean during the 4th Millennium BC", in: Western Anatolia before Troy. Proto-Urbanisation in the 4th Millennium BC? Proceedings of the International Symposium held at the Kunsthistorisches Museum Vienna, Austria, 21–24 November, 2012, OREA 1, B. Horejs and M. Mehofer Eds., Vienna, 2014, pp. 15-41/pp.305-28.

[2] *Arches* http://archesproject.org/

[3] E. Alram-Stern, „Die Ägäische Frühzeit". Band 1: Das Neolithikum in Griechenland. Veröffentlichungen der mykenischen Kommission 16, 1996.

[4] F. Schachermeyr (†), „Sammlung Fritz Schachermeyr: Die neolithische Keramik Thessaliens". Aus dem Nachlaß bearbeitet von Eva Alram-Stern. Veröffentlichungen der Mykenischen Kommission 13, 1991.

[5] TAY Archaeological Settlements of Turkey Project: http://www.tayproject.org/veritabeng.html and Archaeology in Greece Online http://chronique.efa.gr/index.php/fiches/search/.

[6] Historic England http://thesaurus.historicengland.org.uk/ and British Museum thesauri http://www.collectionstrust.org.uk/assets/thesaurus_bmon/Objintro.htm; http://www.collectionstrust.org.uk/assets/thesaurus_bmm/matintro.htm?phpMyAdmin=OYNyINPdn3sQmoXugKH1gcCLSW0.

[7] P. Le Boeuf, M. Doerr, Chr. Emil Ore, St. Stead (current main editors), Definition of the CIDOC Conceptual Reference Model, available from: http://www.cidoc-crm.org/official_release_cidoc.html.

[8] A. Masur, K. May, G. Hiebel and E. Aspöck, "Comparing and mapping archaeological excavation data from different recording systems for integration using ontologies", CHNT Proceedings, Stadtarchäologie, Museen der Stadt Wien (ed.), CHNT Proceedings 2013, 2014. doi: 0xc1aa500e 0x0031b420.