ANALECTA PRAEHISTORICA LEIDENSIA 28

PUBLICATIONS OF THE INSTITUTE OF PREHISTORY
UNIVERSITY OF LEIDEN

INTERFACING THE PAST

COMPUTER APPLICATIONS AND QUANTITATIVE METHODS IN ARCHAEOLOGY CAA95 VOL. I

EDITED BY
HANS KAMERMANS AND KELLY FENNEMA

UNIVERSITY OF LEIDEN 1996

contents

VOLUME I

Archaeometry

Statistics and Classification

Jon Holmen
Espen Uleberg

The National Documentation Project of Norway – the Archaeological sub-project

1 Introduction

The National Documentation Project of Norway is a cooperative project between the Faculties of Art in the Norwegian universities, and is now in its fourth year. The main purpose of the project is to convert information from paper based archives to electronically readable media in order to make the archives more accessible. The project has been working with what can be called the 'collection departments', like the Department of Lexicography, the Department of Folk Music and the university museums with Archaeological and Numismatic collections. The aim is to create a national database for language and culture, where it will be possible to do multidisciplinary studies, combining material from all Norwegian universities.

2 Project organization

The project has its base at the University of Oslo. It uses hand-picked, previously unemployed persons to convert the information (Ore 1995: 278). The workforce is organized in a number of small groups in southern Norway, and four larger groups in northern Norway. The different groups are assigned to different part projects. The people converting the archaeological data do not necessarily have any previous experience with archaeology, but through supervision they are given sufficient education to be able to perform the required text analyses and encoding.

The aim is to create a system that integrates information from several disciplines. Because of this, it is not sufficient to create computerized versions of today's archives. One of the most important aspects when building such a data model, is to have a fruitful dialogue between programmers and professionals in the different disciplines. There is of course no one solution as to how this system should be made, but it is vital that the system is not dramatically different from what is in use today. The cooperating institutes need systems that they feel comfortable with, so that the computerized versions will be of use and will be used by all staff members.

3 Using the data

The resultant information will eventually be more readily accessible to researchers, students, people working with Cultural Resource Management and the general public. Information from the different sections of the project will be combined, so that studies concentrating on a certain area will retrieve information from all the different sources. These sources, (fig. 1) Archaeology, Runes, Old Norse, Modern Norwegian, Dialects, Syntax/semantics, Place names, Folklore and Folk Music will all be connected through the variables Time, Location and Word. This will be accessible for enquiries from Government Planning Agencies, Norwegian Archaeological Authorities, the National Archives, people interested in local history, the Norwegian Mapping Authority, and the Norwegian Language Council. It will be useful in connection with dictionary production, as a writing assisting tool and for primary education.

Combining these sources with an incremental database structure, the system makes it possible to look at an area in a time perspective (fig. 2). Textual information is combined with drawings, photos, maps and sounds to create a Geographical Information System which will eventually include all of Norway. It will be possible to make queries about language development, place names and archaeological sites and finds. The potential inherent in the combination of different sources is especially useful to synthesizing disciplines like archaeology and history.

4 The archaeological sections of the Documentation Project

Norway has five archaeological museums. They are situated in Oslo, Bergen, Trondheim, Tromsø and Stavanger, and with the exception of the latter, all are university museums. Norway does not have a central museum, although the museum in Oslo tends to take a leading role, being situated in the capital. All five museums started as private collections and gradually developed into regional museums. Each museum has a collection of items from its own district. However, previously the geographical division between the museums was not so rigid, resulting in the different museums having artefacts from other museum districts. This means that it is necessary to combine information from all museums to get as complete a picture as possible of the known artefacts.

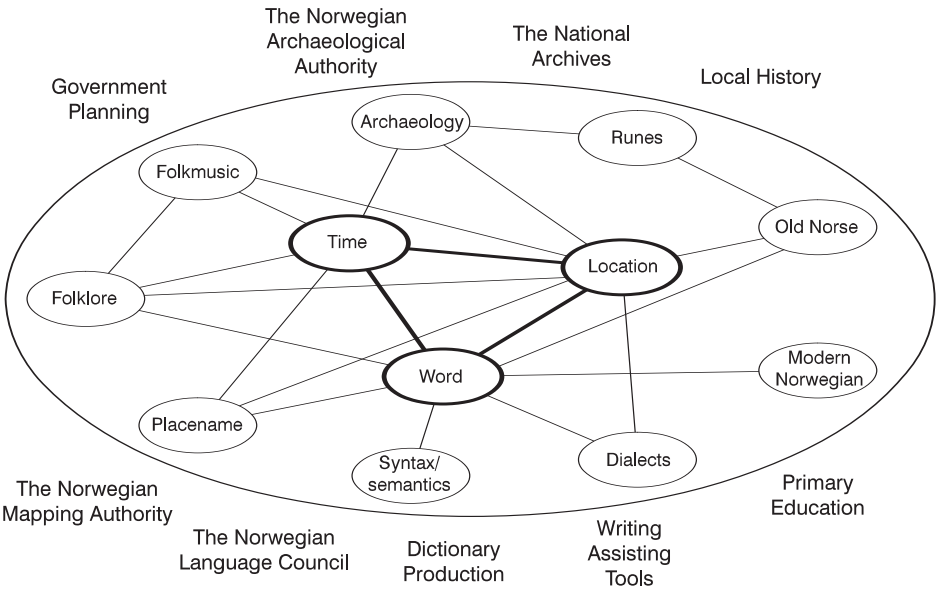

Figure 1. The university information system with its data types, their connections and possible use. (After Ore 1995: 278).

Comprehensive archaeological surveying has been conducted by the Land Use Mapping Agency based at the archaeological museums since 1963 (Larsen 1990: 48). Through this, large parts of the country have already been surveyed. All the resultant information from Oslo university museum, Oldsaksamlingen's district is now stored on computer in a 'free text' database called SIFT (Boaz/ Uleberg 1993: 178-179).

The archaeological sections of the Documentation Project are presently limited to the universities in Oslo, Bergen and Tromsø. Work in Bergen started with converting information on sites, and is continuing with the artefact catalogues. Tromsø has just started, beginning with the artefact catalogues. At the Oslo university museum, the Oldsaksamlingen, the work within the Documentation Project started with conversion of information related to archaeological sites. Since 1993, it has also focused on the artefact catalogues. The artefact catalogues have been converted to machine readable format, and Standard General Markup Language (SGML) is used as a tool to make them more readily accessible.

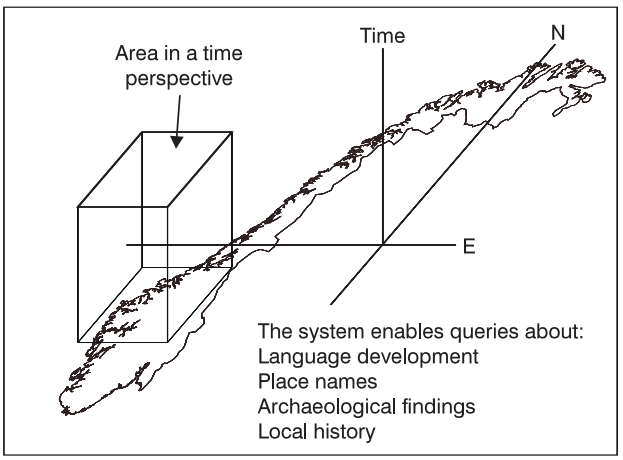

Figure 2. The system makes it possible to look at an area in a time perspective. (After Ore 1995: 281).

5 Ongoing Projects

In addition to the work with the existing archives, the Documentation Project cooperates with ongoing rescue excavation projects. At the moment we are actively collaborating with three projects. One of which has mainly Stone Age excavations, one with Bronze Age/Iron Age excavations, and one is an excavation in a Medieaval town.

Materials from the Stone Age excavations offered the possibility to develop a Geographical Information System to be used on a small scale – to study the artefact spread within one site.

The Bronze/Iron Age excavations gave us a possibility to develop systems to increase the accuracy and efficiency when exposing large areas, and the Medieval town gave

the entanglements of a multilayer site. The methods developed, allows the excavator to have a constant overview over the different structures that are found, both by viewing them on screen, and by printing out distribution plots.

Perhaps the most important aspect is that the data capturing devices that are in use during the excavation make it possible to use the information immediately during the excavation and not just in the final stages of analysis. In addition, the preparation of the final reports can be conducted much more efficiently when the complete data are readily at hand at the end of the field season.

6 The database model

The conceptual model of the database is object oriented, and consists of a number of different relational bases. The concept of an event is a crucial element in the model. An event has been defined as 'something that takes place in time and space, perhaps on account of 'someone' or 'something'' (Rold 1993: 215). There are basically two types of events: internal and external. In this context, all internal events takes place at the museum, external ones out in the field. Internal events include cataloguing, conservation work, etc. Examples of external events are surveying, describing and excavating sites.

The event allows us to make the system incremental, adding a historical depth to the database. It is not a relational database with only the updated information. Every time an action is undertaken, there will be information added to the base. All events will create documents such as artefact catalogues, excavation reports, plan- and profile drawings, photographs, surveying reports and test results. The documents will be in the form of free text, hypertext, bitmap files, scanned documents and pictures. All events will be connected to either an artefact, a site or both. The event makes it possible to retain the information from the original cataloguing as well as incorporating the information from magazine revisions and researcher's special studies on selected artefacts.

All original names are kept when transferring the original artefact descriptions. This means that there will be outdated names on artefacts as well as on places. Lengthy discussions commonly surround the terminology of artefacts. We have avoided this discussion, and will later add standardized artefact names on a higher level, using a meta-language to select for all objects of a particular type, originally given different names in the catalogue. The meta-language will interpret the data without changing the original data (Rold 1993: 218). This will solve problems in cases where old and new terms define the same artefact classes. However, in cases where there is only partial overlap, where one old type now is defined as several new ones, the database will not give precise answers to a query. This situation will gradually be resolved in the future as researchers reclassify these artefacts, and their results are added to the base.

Standard General Markup Language (SGML) is used as a tool to make the converted texts more readily accessible. The SGML is based on formatting a text through adding tags showing the type of information following the tag. In the archaeological artefact catalogues, there are tags for location, material type, artefact type, decoration, dating, and so on. This makes it easier and faster to search in a text, and it also creates a link between a relational database and the free text. In this way, the SGML is a means for structuring a text. One might say that while a database is putting text into tables, SGML is putting a table on a text. The text structure is outlined in a Data Type Description (DTD).

Since a number of different people have written the catalogues at the museum, there are at least as many text structures. A very tight SGML system will give ample opportunity to check that the text is consistent with the DTD, but it will not be possible to incorporate all texts. A system that covers all possible types of text structures, will have become so loose that it is not a structure any more. The final DTD must be somewhere in between these two possibilities.

7 Geographical information

The relation between an artefact and a place needs special consideration in two ways. First, we have to know the present-day equivalent of the old place name, since the original catalogue texts are always used. A problem arises because boundaries between administrative units have changed, and in many cases objects do not have an exact provenance. Because of this, every place name must be associated with a chronological date, indicating what geographical area is covered by that place name.

Secondly, we must decide what to do with artefacts without an exact provenance — perhaps only the parish, county or even just the country is known. One solution is to let the artefact's position be a point somewhere within an area corresponding with the most accurate provenance data. When looking for artefacts from a smaller area, like a group of farms, there is a possibility that artefacts only related to a larger area, like the county, could actually come from that smaller area. Therefore, one must also be able to search for artefacts with a possible provenance within a specified area. This means, that an artefact that cannot be attributed with certainty, should have a geographical location as a point included in a surface area. When users search for artefacts from a certain geographical area, they should obviously retrieve all artefacts where the area of the artefact coincides with or is contained within the search area. In addition, they

should also retrieve the artefacts where the search area is contained within the artefact area, as well as the artefacts whose area intersects with the search area.

8 Conclusion

The National Documentation Project creates a system where databases from different institutions can communicate through the Internet, forming a national database for language and culture. A user will be able to access a client section which will be designed according to the users' needs and use privileges. The databases can be aware of each other, send queries according to a predefined protocol, and interpret the resulting data according to the predefined data models.

The system will allow individual users to create their own interfaces. One such interface can be created in connection with the preparation of an exhibition. Information from different archives related to the exhibited items can be put together in an application running in the exhibition rooms.

There are three major assets of the National Documentation Project of Norway. Firstly, it is an effective system for converting large amounts of data in a relatively short time. Secondly, it promotes a dialogue which is vitally important between system developers and professionals from different disciplines. This ensures the development of effective systems that are sufficiently familiar to be used by all staff members.

A third and final aspect is the increased availability of vast amounts of data. This opens up possibilities, not only for researchers and students, but also for teachers and for the interested public. Most people want to know their local history and what is found at or near their homes. The day is soon at hand when an interested member of the public can turn to the computer in their library to obtain access to information concerning them.

references

Boaz, J.S. E. Uleberg 1993 Gardermoen Project – Use of a GIS system in antiquities registration and research. In: J. Andresen/T. Madsen/I. Scollar (eds), *Computing the Past, Computer Applications and Quantitative Methods in Archaeology CAA92*, 177-182, Aarhus: Aarhus University Press.

Larsen, J.H. 1990 *Om desimering av våre fornminner. Noen resultater av arbeidet med registrering av fornminner for det Økonomiske kartverket i 1980-årene*. Universitetets Oldsaksamlings Årbok 1989/1990:47-60.

Ore, C.E. 1995 Making an Information System for the Humanities, *Computers and the Humanities* 28, 277-282.

Rold, L. 1993 Syntheses in object oriented analyses. In: J. Andresen/T. Madsen/I. Scollar (eds), *Computing the Past, Computer Applications and Quantitative Methods in Archaeology CAA92*, 213-220, Aarhus: Aarhus University Press.

Jon Holmen and Espen Uleberg
The Documentation Project
Faculty of Arts
University of Oslo
P.O. Box 1102
Blindern
0317 Oslo
Norway
e-mail: jon.holmen@dokpro.uio.no
espen.uleberg@iakn.uio.no