Reference Collections
Foundation for Future Archaeology

Edited by A.G. Lange

Colofon

Reference Collections
Foundation for Future Archaeology

Proceedings of the international conference on
the European electronic Reference Collection
May12-13, 2004,
ROB, Amersfoort, The Netherlands

A.G. Lange (editor)
English corrections and technical editing
Drs. Kelly Fennema, Leiden
Cover design
Drs. Menne Kosian, ROB
Original design
Bert van As, ROB
Design, layout and prepress
Elvé Ex Press, Den Haag
Printing
Plantijn Casparie, Almere

ISBN 90-5799-059-8

Rijksdienst voor het Oudheidkundig Bodemonderzoek

SENTER

NWO
Netherlands Organisation for Scientific Research
Humanities

Rijksdienst voor het Oudheidkundig Bodemonderzoek
State Service for Archaeological Heritage
PO Box 1600
3800 BP Amersfoort
The Netherlands

2 Classification and archaeological knowledge bases

Torsten Madsen

Abstract

Reference collections are classifications in action and electronic versions imply the use of databases to store and organize the information. The way classifications are implemented in archaeological databases is far from unproblematic, however. This paper will initially summarize the role of classifications in archaeology, followed by an outline of the problems that may arise when using classifications in connection with databases. Finally it tries to clarify what this may mean to electronic reference collections.

1 Categories and classes

'Why do we have to learn all these types by heart?' students frequently ask me. 'We would rather learn more about the past!' For many years I did not give a consistent answer to that question. Rather when teaching typology and classification as a theoretical and methodological issue, I focused on the nature of types and classes. How are they structured, what is the proper way of creating them, are they artificial constructions on our part or do they reflect 'real' types, recognized in the past as well? All these are standard questions discussed intensively in the literature over the years.
Reading the book 'Archaeological typology and practical reality' by Adams & Adams (1991) changed my perspective somewhat. With one of the authors being a philosopher the book goes much deeper than ordinary texts on classification and typology in archaeology, and it shows systematically how classifications and typologies are closely related to basic human observation and categorization of the world around us. Only with typologies, which in their definition includes a demand for logically constructed, unambiguous sorting keys, do we have a categorization principle that separates itself from what ordinarily takes place in human society.

Categorization is our way to meaningfully structure the world we observe, and it forms the basis for efficient communication when we share the structure and its meaning with others in a group. Categorization is a basic means by which human beings create culture. In archaeology, when we insist that students learn classifications by heart, we are in fact integrating them into the archaeological subculture, setting them up with the means to communicate meaningfully with their peers, and at the same time we indoctrinate them with a meaningful structure, through which they can understand the past.

The degree of detail to which we take categorization depends on our needs. To me a camel is a camel, but Bedouins have around one hundred categories of camels. To me there is wet snow and dry snow, but to the Inuit there are some fifty categories of snow. In archaeology there is a great variation in the degree of detail as well, but here variation occurs between very small groups of people, even down to individuals. I recall a colleague at a local museum in Denmark who was setting up a recording system for the museum's excavations. He had some fifty categories for Iron Age pottery sherds but only one category for flint. My own recording system is not very different as I have only one category for Iron Age sherds and numerous categories for flint debris and artefacts.

The way changes in categorization of the world take place in society is complex and apparently unorganized. Obviously the forming and subsequent influence of subcultures plays a major role, but exactly how and why particular views become dominant is an interesting sociological problem. In archaeology, being a research discipline, we can expect changes in the way we categorize things to be clearer and not least more deliberate and controlled. Yet, within archaeology we can also find groups and alliances that use specific ways of categorizing their observations in a struggle for dominance and acceptance, but also of course to find ways of creating a better and more relevant understanding of the past. The nature of classes also varies from subject area to subject area, from country to country, and not least it has changed with time.

2 Traditional classification

From the outset archaeological classifications have been restricted to complete objects and not to parts of objects. It seems natural to think in complete objects even if details of the objects constitute the foundation of the classifications. Such is the preoccupation with complete objects that a type is formally 'defined' through a typical example to which you attach an often informal description based on details of the object. Since no formal definition of typehood exists, it is imperative that you have a typical object or a picture of it in front of you to learn what constitutes the type. This is the exact same way that children learn the categories of the world they have entered. You can point to a dog in the street and say 'dog', or you can point to a picture of a dog in a book and say 'dog'. Thus reference collections are a must when dealing with this type of classification. They can either be collections of real objects or collections of pictures of objects. When teaching students we use both, the same way as we were taught way back when.
It is important to note this kind of classification creates very fuzzy types. As a student I took part in a seminar organized by a teacher who was very interested in these problems. We took all the axes of the Single Grave culture in the museum and classified them according to the classification set up by P.V. Glob (1944). Most of these axes had been used as the basis for his study as well, and in the back of the book he had provided lists of inventory numbers sorted according to his classification. To our surprise and dismay, we found that agreement between his assignment of axes to classes and ours was often minimal.

3 Analytical classification

In the middle of the last century an interest arose in formal classification of descriptive elements. We normally refer to this as analytical classification following Rouse, who apparently was the first to apply it (Rouse 1939). In Europe in the 1960s, Jean Claude Gardin (1958, 1967) was a key figure together with Mats P. Malmer (1962, 1963), and the procedure has been given a central role in education in Denmark since the late 1960s.
The basic idea is to formally define descriptive elements of artefacts in such a way that the elements can be either measured, be present in some state, or be absent. Classification is then a question of setting up a logical sorting system for objects specifying, mostly in a hierarchical way, particular ranges of measurements and/or the presence of particular elements/element states for achieving typehood.
There is a significant difference here with the way classification had worked till then, which is that you do not need a reference collection in order to classify an object. It is simply a question of taking particular measurements and recording the presence and state of particular elements. All you need is the sorting key and then of course a clear understanding of how the elements are defined and measured, which in many cases means that you need a sort of reference collection for the analytical classification, but not necessarily one that shows particular objects.
Most major classification systems created since the beginning of the 1970s in Denmark use this approach to classification. They produce clearly defined monothetic type divisions, and although these are in fact 'artificial' classes, using them in further analyses has produced amazingly good results (see Ilkjær 1990 for an example on Iron Age spearheads).
Over the last 15 years a procedure has been developed, in which analyses are carried out directly on the analytical classifications. The idea for doing this dates back to the 1960s, but was not an attractive approach before personal computers became readily available. It is primarily Correspondence Analysis that is used, and mainly highly fragmented materials and richly decorated materials have attracted this approach. The results have been very, very good and the message is beginning to register: there is no reason for establishing detailed object classifications in the first place; it is sufficient and even better to work directly with detailed analytical classifications of descriptive elements (see Madsen 1988 for a series of examples instrumental in forming this approach).

It is still a question of classification though, and we still need some sort of reference collection for these classifications, but they are of a different kind. We are at a level of detail, where no studies will ever use the exact same set of classes, and the number of classes may be truly staggering. A colleague of mine working with animal style decorations on brooches from Early Medieval time operates with a total of close to two thousand classes and modification occurs all the time (Karen Høilund Nielsen, personal communication). So a 'reference collection' here is more of a documentation of classes used in a particular study.

4 Dynamics in research

I am certain that in the future we will see many more studies based exclusively on analytical classifications. As a result we will also find that object classifications will play a diminishing role. In principle it makes no difference to the research process, however. Classification, whether of objects or elements, is not an end in itself, even if it is a way to structure and communicate our observations. The meaningfulness of our classes will have to be tested against the contextual information that is also part of our world of observation.
The bottom line is that classification is an operational tool in archaeology used to create meaningful structures from observations. These are then used in analysis, and as a basis of internal scholarly communication. Our successes as researchers depend to a great deal on our ability to interact with our classifications and to dynamically change these to create new meaning that better explains our observations in terms of past culture. Any classification considered to be the ultimate solution in a humanities discipline is a testimony to fossilization of research. The purpose of standardization is to keep things unchangeable and controllable, which is a great thing in administration and industrial production, but not in research into human culture, where the aim is to continuously construct meaning and insight into our life as human beings.

5 Fossilized knowledge

This leads to the problem of classifications in large, permanent archaeological knowledge bases. Classifications and typologies published in print are historic statements that may immediately be negotiated and modified by others in new publications thus creating a dynamic development. Databases in the form of huge knowledge bases designed to last for decades at least and meant to be loaded cooperatively with information over time poses a real problem here, because it is not obvious how the classifications and typologies embedded in them can be negotiated and modified, and indeed if technically possible by whom they should be negotiated and modified. All experiences so far show that the classificatory base of such databases fossilizes from the very beginning, and thus becomes counterproductive to research.
Fossilization in a knowledge base primarily takes place because the classification system used is not up to date, it is not reflecting current usage and current trends in research. As the classification system is considered authoritative and an absolute key for searching the knowledge base it cannot be changed, however, without a complete reclassification of the content of the knowledge base. Neolithic causewayed enclosures were unknown in Denmark when the classification scheme for the Danish SMR was created. Today we know of at least 50 of these monuments, but in the SMR they are simply recorded as settlements, and there is no way to separate them through a search.

With fossilization comes corruption of the meaning of the classification system, because even if the classes of the knowledge base do not change, our understanding and interpretation of them do, and at the same time new categories turn up that somehow will have to be squeezed into the existing system, thus changing the de facto meaning of the existing classes. Again, when the Danish SMR was set up, no categories were created that could represent the findings of two or three pieces of flint debris found in a field during survey, because

the whole classification system was contemplated in cultural terms. This kind of 'background noise' is today more often than not classified as 'settlement' in the absence of a more suitable category, and it thus completely corrupts the meaning of settlement in the knowledge base.

6 Object oriented design

We certainly need large databases to store the vast quantities of information we produce, to secure this information, and not least to disseminate it through networks. We should be very careful however when deciding what to store, and how to store it. It is important that we make a distinction between what can be considered historical data and what is current ad hoc interpretative information. Historical data should be stored exactly as created. By historical data I mean for instance to the results of an excavation (the excavation report), recording of a site at a given point in time, an illustration of an object together with its administrative identification, or a classification presented by a particular person at a specific occasion. The interpretative information, which certainly includes the way we currently tend to classify the historical data, should be negotiable at all times.
We can store all the historical data as simple documents (an excavation report as printed, for instance), but that is certainly not a satisfactory solution, especially not as some of the newer information stems from well-structured databases. On the other hand, we cannot allow the storing in a common database format of excerpts or interpreted versions, as this would breach the integrity of the historical information. The solution is to use object oriented designs combined with metadata descriptions. It is thus fully possible to store very different database structures with different content definitions in the same physical structure without loss of information.

The problem with this approach, however, is that it becomes very difficult to retrieve information from the knowledge base across elements in the database that differs in logical structure and content definition. We lack the overall authoritative classification of the content, but to include that is, in my opinion, out of the question. We are therefore left with only two possibilities, both of which should be used. One is the use of search engines of the type known from the Internet, with all the weaknesses they have. The other is an interface to the knowledge base that allows researchers individually or in groups, publicly or privately to set up and maintain formal classifications to the content of the knowledge base. In this way multi-vocality in interpreting the content of the database can be established. Technically all of this could be set up today, but apart from a few notable uses of object oriented design, archaeology is a far, far way from achieving this kind of solution (for instance ArchaeoInfo and Intrasis[1]).

7 Reference collections

We can distinguish two types of traditional reference collections. One consists of proper physical collections, the other of printed type inventories as part of ordinary scientific publications or as regular guidebooks on particular categories of material.

1 www.archaeoinfo.dk / www.raa.se/uv/intrasis

In order to be of practical use, the physical collection will have to be ordered according to one particular classification system, in the same way as books in a library are placed on shelves according to some classificatory principle. The classification scheme is thus an integrated part of the traditional reference collection by means of the physical organisation of the collection. To have one physical collection representing more than one classification scheme would be very unpractical. Physical reference collections will also have to exist undivided at one location, as they will be of little use otherwise.
Building a reference collection at a specific location, say a museum, may be difficult because not all objects covering all the types of a specific classification may be found at that museum. For instance it has not been possible to establish a complete reference collection of Neolithic flint axes for teaching purposes at my own institution, even though we are associated with the second largest museum in Denmark. Physical reference collections thus tend to be restricted to more or less technical aspects on a common level of occurrence, and they seem to be relatively rare. In Denmark they only exist for teaching purposes in connection with the universities.

The printed type inventories in practice share one of the limitations of the physical collections. Due to the ordered, sequential presentation they are bound to present one particular classification system, and although it would be possible to present alternative classifications with cross references to the illustrations, this is very seldom done, partly because it is not the intention of the publications to do so, and partly because, as with the physical collections, it would be impractical to deal with. The great strength of the published type inventories is of course that they, in contrast with the physical collections, are not limited by the location of the objects, both in terms of what can be brought together in an inventory, and of the location of use. For this reason they are much more practical to work with than the physical collections, and in Denmark for instance they form the backbone of all referencing of objects in excavation reports and publications.
An electronic reference collection may be viewed as a knowledge base containing two sets of quite different information. One is information about concrete objects including ideally various illustrations of the objects, description of the find circumstances associated with them and administrative information on their life story, including their whereabouts, their physical condition/treatment, and references to where they have been published and discussed in the literature. This is all proper descriptive information of a historical nature.
The other is a documentation of the description and classification system to be applied to objects. This is in fact also information of a historical nature, as all description and classification systems in principle are mental constructs created at some point in time by specific persons. However, they are on a totally different level than the objects as they can exist, in principle at least, without the objects.
The combination of these two sets of information in reality constitutes the electronic reference collection. Electronic reference collections thus simply consist of a formal linkage between two sets of information in a knowledge base: historical information on objects and historical information on our mental constructs of description and classification.

8 Benefits of electronic reference collections

As with the printed type inventories, the electronic reference collection lacks the possibility of physical contact with the objects, but shares all the benefits of the former and has a few potentials of its own. Most importantly the objects of the collection consist of virtual representations and are thus not limited to a particular physical order. The order of objects is determined by the way they are linked to a classification system, and as any object may be linked to an unlimited number of classification systems, they may be part of many different orderings at the same time. It also allows for comparisons and cross-referencing of different classification systems applied to the same material.
An electronic reference collection of this kind would be extremely useful to research. The simultaneous application of different classifications to the same material in a regular knowledge base makes it possible to compare which description and classification systems have been applied to which objects, by whom and when. Very often you will have to dig deep in order to understand the implications of a particular classification in the literature, and in most classification studies you will find an in-depth analysis of previous work in that particular field of study.
An electronic reference collection could also (and easily) be given the same structure as a printed one, but that would lead directly to fossilization as with all other knowledge bases that embed the interpretative statements with the basic information. There is one fundamental difference between electronic reference collections and printed type inventories, and that is that the latter has a publishing date, which is taken into consideration when you evaluate and use its content. The electronic reference collection will always appear as current regardless of whether its content is old and outdated or not.

9 Conclusion

Classification is an active research tool through which we create structure and assign meaning to our observations. Classifications are entirely our constructs, and the meaning of classes is our meaning, irrespective of whether comparable categories existed in the past or not. By creating new classifications and restructuring old ones we create new meaning and new knowledge, and the constructions we create become important elements in the scholarly debate. Electronic reference collections can become a very valuable element in this process, but a large database, as it will be, tends to fossilize, its content can easily be controlled by a few and there is a danger that it becomes authoritative in its own right, no matter what it contains. If this happens, the reference collection could become a hindrance rather than a help to the research process. Therefore we must be very careful when we set out to design an electronic reference collection.

References

Adams, William Y. and Ernest W. Adams, 1991. Archaeological typology and practical reality. A dialectical approach to artefact classification and sorting. Cambrindge: Cambridge University Press.

Gardin, Jean-Claude, 1958. Four codes for the Description of Artifacts: An essay in Archaeological Technique and Theory. American Anthropologist 50, 335-57.

Gardin, Jean-Claude, 1967. Methods for the descriptive analysis of Archaeological Material. American Antiquity 32(1), 13-30.

Glob, P.V., 1944. Studier over den jysk Enkeltgravskultur. Aarbøger for Nordisk Oldkyndighed og Historie.

Ilkjær, Jørgen, 1990. Illerup Ådal 1. Die Lanzen und Speere. Jutland Archaeological Society Publications XXV,1.

Madsen, Torsten (ed.), 1988. Multivariate Archaeology. Numerical Approaches in Scandinavian Archaeology. Jutland Archaeological Society Publications XXI.

Malmer, Mats P., 1962. Jungneolithische Studien. Acta Archaeologica Lundensia, Series in 80, No. 2.

Malmer, Mats P., 1963. Metodproblem inom järnålderens konsthistoria. Acta Archaeologica Lundensia, Series in 80, No. 3.

Rouse, Irwing, 1939. Prehistory in Haiti, a Study in Method. Yale University Publications in Anthropology 21.