Toward standardised vocabularies for Norwegian archaeology

Experiences at the Museum of Cultural History, University of Oslo

Espen ULEBERG, Museum of Cultural History, University of Oslo, Norway
Mieko MATSUMOTO, Museum of Cultural History, University of Oslo, Norway
Alexis PANTOS, Museum of Cultural History, University of Oslo, Norway
Letizia BONELLI, Museum of Cultural History, University of Oslo, Norway

The Norwegian university museums have worked over several years to create common vocabularies for use in their National database system. The first online version of the artefact database was based on SGML-tagged (Standard Generalized Markup Language) catalogue texts. The tagging schema was defined for the archaeological catalogues, and contained tags for artefact name, material type, location etc (Holmen and Uleberg, 1996). This made it possible to query the original texts, but met with challenges created by changes in Norwegian orthography and archaeological terminology that took place over the more than 150 years of archaeological catalogue texts. A first step towards standardisation was made when the SGML-tagged texts where imported into the database with the spelling of some words in the older text changed according to the new spelling rules that were applied from 1917 (for example the word sword that changed from "sværd" to "sverd").

The archaeologists at the five archaeological museums in Norway have always had close contact and attempted to maintain close terminology. One example is the guidelines for describing struck stone artefacts that archaeologists from the museums in Tromsø, Bergen and Oslo compiled together (Helskog et al. 1976). The new artefact database created an opportunity to continue to work on terminology standards. The aim was to create a common standard terminology for periods, artefacts, find categories, materials and find circumstances. The work started at the museum in Oslo and later expanded to include all the museums. The terminologies used across the museums are now largely the same, but with varying find categories reflecting the variation in assemblages in different parts of Norway. The period terminology has been sent to PeriodO (Perio.do), and artefact terminology and find categories are mapped to Getty's AAT as part of the ARIADNEplus project.

Most of the terminology can be mapped directly, but some words can only be mapped to more general terms. Different research traditions have also made it difficult to agree on a single set of terms. At present, each museum can supply terms to their own lists, in order to describe artefact categories that are missing from the original terminology. There are also a number of artefacts that are still described with original terminology, while it may be tempting to automatically update these terms, in most cases it is necessary to study each artefact individually to determine the correct modern typology. It should also be mentioned that so far the terminology has only been made for one of the official languages in Norway.

Another step toward standardisation was made in 2009 when a fixed list of Norwegian cadastral entities was added to the artefact database. The table contains the units county, municipality and farm, and the table combines cadastral lists from 1886, 1950, 2000, and later administrative changes in Norway. The historic data helped to relate find information in early catalogue texts to present administrative units. The move toward standardisation continued in 2010 with the adoption by Norwegian university museums of the Swedish system Intrasis (intrasis.com). The Intrasis software allows each field leader flexibility in the design of the documentation of each excavation. In order to facilitate combination of data from separate excavations and integrate them with the existing artefact and photo databases the Norwegian museums opted for a stricter implementation of the software (though still allowing for the inevitable exceptions on some excavations). As part of the national cooperation, a working group described how Intrasis should be used, and these guidelines are updated every year.

This shared practice was the backdrop for the recent infrastructure project ADED (Achaeological Digital excavation Documentation) that combines detailed excavation documentation in Norway and makes it accessible online. The find categories used in ADED and the artefact database also relates to the find categories used in the Norwegian HER (Heritage and Environment Register), Askeladden. Work is recently done to map this terminology to the terminology used in the Danish equivalent Fund og Fortidsminder (https://www.kulturarv.dk/fundogfortidsminder) and to Getty's AAT.

More recently the Museum of Cultural History (MCH) has expanded its use of 3D-documentation of excavations and artefacts. The 3D archive and dissemination infrastructure project BItFROST (https://www.khm.uio.no/english/research/projects/bitfrost/index.html) works to increase the use of 3D documentation in research and teaching. Standards for sustainable storage and dissemination of 3D models are still very much in flux, and BItFROST aims to better define the structures at MCH and to facilitate accessibility and migration to newer systems and data formats in the future. BItFROST aims to support open source technologies as an approach for sustainable development, and is built around the 3DHOP viewer for dissemination and shares resources and strategies with the Dynamic Collections project in Lund, Sweden.

Standardisation through mapping to international standards like Getty's AAT is good, but the specificity and complexity of local cultural material traditions mean that such common word lists will always have their limits. One should also consider mapping of terminology between neighbouring countries as different environmental and historical backgrounds will be reflected in the use and sustainability of the vocabulary. An example is the Nordic cooperation "Nomina Rerum Mediaevalium" that published drawings and terminology for mediaeval artefacts in the five Nordic languages (Norwegian, Danish, Swedish, Icelandic and Finnish). Even though precise translations / mappings can be more easily found between these neighbouring countries, there will always be undeniable differences of vocabularies to describe events and phenomena that have no universal term.

The artefact data are published in original language under unimus.no; as downloadable datasets, API and at the Unimusportal (unimus.no/portal). The 3D-models in BItFROST and the excavation data from ADED will be published at separate webpages under unmus.no. IDs from Askeladden and the databases for artefacts and images are used to provide inks between these instances (Fig. 1). Future developments will include wider linking and sharing of data.

Linking at site level:

Site links

Askeladden	SiteID
Artefact database	SiteID ArtefactID
Photo database	SiteID ArtefactID ImageID
ADED	SiteID ArtefactID ImageID

C238 Rattle from Vågå. Photo: Ulla Schildt, KHM CC 4.0 BY SA

The aggregation and standardisation of archaeological metadata in Norway has prepared the ground for sharing data within ARIADNEplus and has given researchers and the general public better access. One should however be aware of the distancing effect that technology can have, and that digital data creates different 'data-imaginaries' (Huggett, 2022) determining our understanding of past and present. As such we should be conscious of the limitations highly regulated vocabularies can have on our understanding of material culture. That said, there is no doubt that increased standardisation will facilitate communication and collaboration across diverse archaeological and heritage specialisms. An important challenge is to have more institutions and nations participating in shared standards. Only this can give access to data across national borders and between institutions in meaningful ways.

Helskog, K., Indrelid, S. and Mikkelsen, E. (1976). 'Morfologisk klassifisering av slåtte steinartefakter', Universitetets Oldsaksamlings årbok. 1972-1974. pp. 9 – 40. Oslo.

Holmen, J. and Uleberg, E. (1996). 'Getting the most out of it – SGML-encoding of archaeological texts'. Paper presented at CAA 1996, Iasi, Romania. Available at https://www.dokpro.uio.no/engelsk/text/getting_most_out_of_it.html (Accessed 18.07.2022).

Hugget, J. (2022). 'Data legacies, epistemic anxieties, and digital imaginaries in archaeology', Digital, 2(2), pp. 267-295. DOI: 10.3390/digital2020016