

DATABASES AND STANDARDISATION – EXPERIENCES FROM ATHENS

This project (*Archive and database of Swedish archaeological research in Greece*¹) is based at the Swedish Institute at Athens and aims to make available the information and material from all Swedish excavations in Greece from 1894 onwards. By organizing and complementing the Institute archive and by creating an online database comprising digitized excavation documentation and information about the archaeological material, we want to make the material more accessible to researchers. Our paper discussed views and experiences concerning databases and standardisations that emanates from the work in this project.

Archaeological material is typically catalogued in structurally simple databases comprising information about objects, find locations, dating and publication data. The fields and field values collected tend to be superficially similar regardless of who designed the various databases, and regardless of when they were designed and what tools were used. However, different excavations and different individuals use different conventions, and the resulting databases may in reality not be all that similar when you look at individual fields. There will be differences in the structure of geographical information, dating, object type and material specifications and in other data, making a detailed and exact comparison difficult.

There are several initiatives underway, nationally and internationally, to establish standards for data formats. However, not only will these take time and substantial effort to implement on a larger scale, but they will also not be very helpful when it comes to existing databases, collected over many decades or even centuries of excavation. It is of course possible to change the structure of a database, but there is a risk that this will destroy information, since fields in the original database may not have an exact correspondence to any field in the chosen standard. Consequently, consolidating multiple databases in a single location, in order to enable searches across multiple databases, is not an easy task if you also want to preserve the original information unaltered.

We presented a method of achieving powerful consolidated search across multiple disparate archaeological databases, by introducing an additional metadata set, to which a loose mapping can be made from the original databases. The search index will contain the original data plus the common

¹ The project is funded by the Swedish Foundation for Humanities and Social Sciences (Riksbankens Jubileumsfond).

metadata, enabling simple search across all databases, as well as more specific searches in a single database for anyone who is familiar with the structure.

The basis for this solution is a document-based search index such as Apache Solr, and we will use a set of Dublin Core (“DC”) metadata fields for the common mapping. These DC fields typically define information on an abstract level, and multiple fields from the original database will be mapped to individual DC fields.

Deciding on a simple metadata set like this can enable us to implement “metasearch” engines that do searches over multiple databases belonging to different institutions on a global scale, and even meaningful searches over multiple disciplines. By ensuring that every record has a well-defined “home” and a globally unique ID, this will make it easy for researchers to find individual items as well as correlations, on a scale hitherto impossible. Data can still be maintained and archived where it belongs, and updates of the search index can be performed in real-time or on demand.

Along with the data search mechanism, we are also planning an image-based search and presentation mechanism, using polygon-based annotations of image files to add searchable text and geographical coordinates. We also envision mechanisms by which visitors can do limited annotation of data and images, with an approval workflow to enable retention of control by the owner of the data.

HANS LISS, ANNE INGVARSSON-SUNDSTRÖM

Swedish Institute at Athens

Hans@Liss.nu, anne.ingvarsson_sundstrom@gustavianum.uu.se