

The lifecycle of pottery data

A case study in archival practice

Nicole HIGH-STESKAL, Department for Image Sciences, Donau Universität Krems, Austria

Laura REMBART, Austrian Archaeological Institute of the Austrian Academy of Sciences, Austria

Keywords: pottery, legacy data, archiving, archaeology, accessibility

A recent project in Aswan, Egypt, resulted in the discovery of an exceptionally large number of pottery fragments. While the large quantity of pottery in combination with the stratified contexts was seminal to the study authors in gaining more detailed insights into the commercial and social life of the inhabitants of Hellenistic, Roman and late antique Syene/Elephantine than expected, it also resulted in considerable problems regarding the publication of the data and the management of such large amounts of information.¹ Throughout the project, the pottery had been processed according to a consistent workflow that included the quantitative assessment of all pieces and the detailed documentation of all diagnostic pieces which were entered in an MS-Access database that was originally created for the documentation of pottery in Ephesos in 1998 and had since been slightly adapted.² In order to publish the entire database of 16,000 diagnostic sherds it was necessary to seek out new and digital forms of publication since the number exceeded what could sensibly be published in a traditional format. Furthermore, the database had been created in an outdated mdb-format and it was necessary to secure the data in its entirety.

Through the FWF-funded Open Research Data project “Wares, Types and Fabrics. The Upper Egyptian Contribution to the LCP”³ the pottery data collected in the MS-Access database was entered into the open access pottery database of the Levantine Ceramics Database (LCP; <https://www.levantineceramics.org/>). The goal of this project had been to upload the original database to the LCP, however, several issues arose in the upload process and necessitated a re-evaluation of the data lifecycle of pottery data and an adjustment of the online publication strategy in order to remedy the issues. Instead of uploading the pottery data to the LCP in its entirety, only the most important pieces were uploaded (ca. 2,500 entries) along with entries on all wares and petro-fabrics, in addition to accompanying photographs, thin section photographs, and drawings. Additionally, the entire original MS-Access database was exported to Excel, cleaned with OpenRefine, enhanced with additional information and cross-linked with the corresponding information in the LCP and other data to increase both their findability and interoperability. These lists represent the complete pottery record and are currently in the process of being curated and prepared in a sustainable data format for upload to ARCHE, the digital archive of the Austrian Academy of Sciences.

As a result of this experience, a guideline for working with pottery data at the Austrian Archaeological Institute was created so that future projects can avoid detours made in earlier projects and support colleagues in developing data management plans before they begin their projects.⁴ The following four rules adapted from the FAIR principles govern the guidelines:

1. Document everything (individuals involved, publications and talks connected to the data, short descriptions of steps taken in workflow, create concordance lists,...) so that the metadata fields can be created as precisely as possible.
2. Be consistent and transparent (terminology, language, file types,...) so that the data is not marred by errors.
3. Contextualize data (link to chronology tools, gazetteers, other pottery databases,...) so that the data becomes findable to other people and also machines.

¹ See also High-Steskal et al. 2017. On the study of the pottery see: Rembart 2018; Katzjäger 2017; Peloschek 2015; for a full list of publications on the site, see High-Steskal 2019a.

² Kerschner et al. 1998.

³ FWF ORD 69; project led by Sabine Ladstätter; based on FWF project P 23866 “Housing in Antiquity in Syene/Elephantine, Upper Egypt.”

⁴ For the guideline, see High-Steskal 2019b.

4. Make data reusable (what information do users need to understand dataset? Can data be downloaded? Have clear standard licenses been formulated?). Reuse increases the value of data and the importance of the scholar who created it.

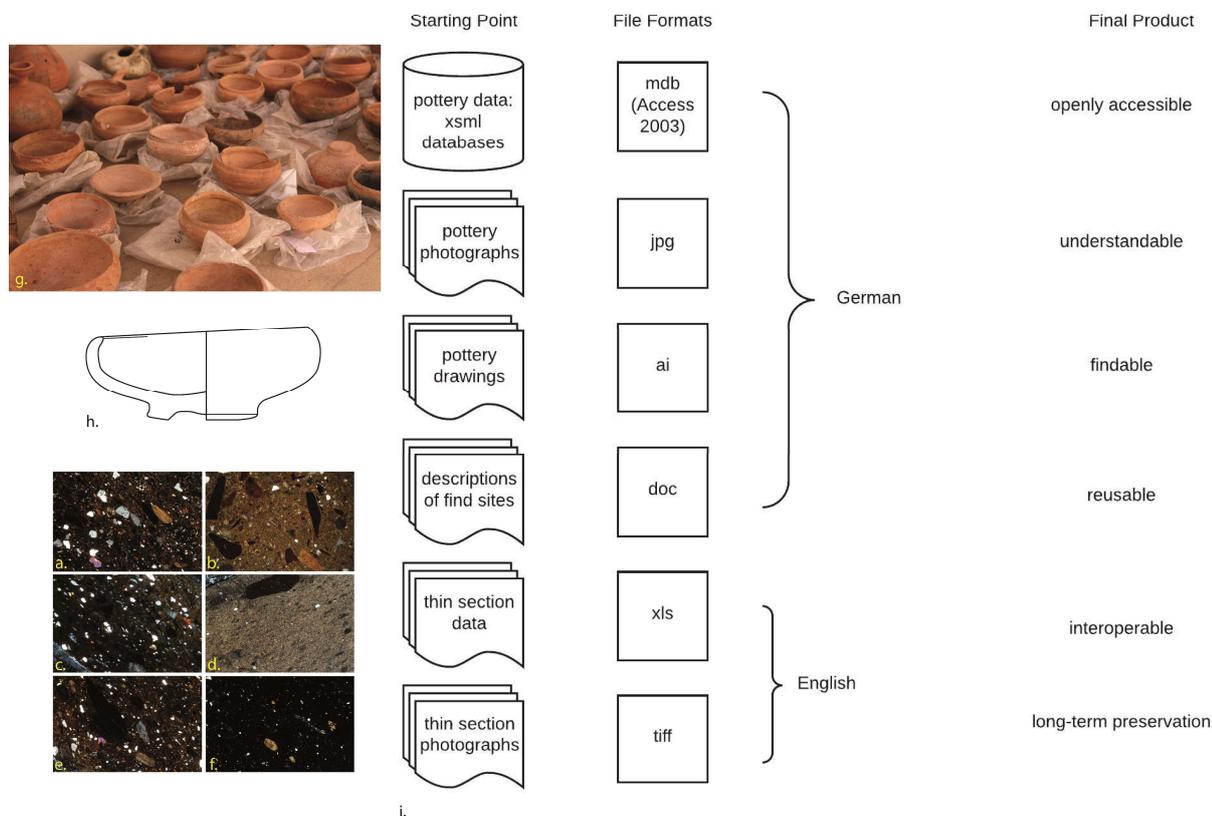


Fig. 1.a-f : thin section samples of pottery from Aswan (© OeAI-OeAW/Lisa Peloschek); g: pottery from Aswan (© OeAI-OeAW/Niki Gail); h: pottery drawing of Aswan material (© OeAI-OeAW/Laura Rembart); i: data types and formats created while processing pottery from excavations in Aswan (© OeAI-OeAW/Nicole High-Steskal)

During the creation of the guideline a conscious decision was made against using technical jargon and in favor of formulating the necessary steps in as much detail as possible while highlighting the benefits of curating data for long-term preservation. The reason for the level of detail in the guideline was to help all researchers, especially the less confident computer users.

In addition to creating more awareness for data management and the need for archiving research data, it was also necessary to engage in broader discussions on the data life cycle with colleagues. To a certain extent curation and archiving processes still hold an uneasy place within our field because they have not yet been accepted as a serious form of scholarly output.⁵ This is expressed, for example, in the inability to list archived datasets in departmental databases that monitor scientific output or also the lack of acceptance for listing most forms of digital contributions in official CVs.⁶

Encouraging archival practices among scholars more broadly requires a multi-pronged approach: educating all individuals on methods and issues in archiving their digital data; including data stewardship as an important aspect of the research process; and accepting archived datasets as valid scientific output of researchers.

⁵ Sobotkova 2018, 117.

⁶ The American Institute of Archaeology recently published the following recommendations under the heading 'Guidelines for the evaluation of digital technology and scholarship in archaeology' here: <https://www.archaeological.org/programs/professionals/careers/tenure/> (accessed August 21, 2019).

References

- High-Steskal, N., Rembart, L., Schwaiger, H. and Ladstätter, S. (2017). Wares, Types and Fabrics. The Upper Egyptian Contribution to the LCP (digital humanities Austria 2017, Innsbruck on December 2017).
<http://doi.org/10.5281/zenodo.3373303> (Accessed: 09 September 2019)
- High-Steskal, N. (2019a). 'Private Houses in Syene and Elephantine and the Paternouthis Archive.' *Jahreshefte des Österreichischen Archäologischen Institutes in Wien*, 87, pp. 123-143.
- High-Steskal, N. (2019b). Anleitung zur Archivierung von Keramikdaten (Version 1). Zenodo.
<http://doi.org/10.5281/zenodo.3385579> (Accessed: 09 September 2019)
- Katzjäger D. (2017). *Spätantike Alltagskultur auf Elephantine, Oberägypten. Typochronologie, Funktion und kulturelle Diversität der keramischen Funde aus den Wohnhäusern um den Chnumtempel*. Dr. Paris Lodron Universität Salzburg.
- Kerschner, M., Plattner, G. A. and Ladstätter, S. (1998) 'Das digitale Dokumentationssystem von Keramikfunden der Grabung Ephesos.' *Jahreshefte des Österreichischen Archäologischen Institutes in Wien*, 68, pp. Beibl. 47-62.
- Peloschek, L. (2015). *Cultural Transfers in Aswan (Upper Egypt). Petrographic Evidence for Ceramic Production and Exchange from the Ptolemaic to the Late Antique Period*. Dr. Universität für angewandte Kunst Wien.
- Rembart, R. (2018). *Kulturelle Identitäten Oberägyptens in ptolemäisch-römischer Zeit am Beispiel von Syene. Typochronologie, Akkulturation und wirtschaftsarchäologische Interpretation von ausgewählten keramischen Fundkomplexen aus der Stadtgrabung*. Dr. Paris Lodron Universität Salzburg.
- Sobotkova, A. (2018). 'Sociotechnical Obstacles to Archaeological Data Reuse.' *Advances in Archaeological Practice* 6, 02, pp. 117–124. <http://doi.org/10.1017/aap.2017.37> (Accessed: 09 September 2019)