The ARIADNEplus integration of archaeological datasets

Franco NICCOLUCCI, PIN, Italy[1]

Keywords: Archaeological datasets, integration, standards

The ARIADNE and ARIADNEplus projects

ARIADNEplus builds on the success of ARIADNE (Archaeological Research Infrastructure for Archaeological Data Networking in Europe, http://www.ariadne-infrastructure.eu/), a project funded by the EC in FP7 for the period 2013-2017 (Niccolucci and Richards, 2013; Meghini et al., 2017). ARIADNE succeeded in integrating archaeological datasets in its Registry, with about 2,000,000 datasets recorded and managed in the ARIADNE Catalogue, according to the FAIR principles; building a community of use consisting of about 11,000 archaeologists engaged with its services, corresponding to one third of all European archaeologists and probably more than 50% of those using some computer support in their research, with an even higher percentage if only early career researchers are considered; proposing standards for the sector as CRMarcheo, an extension of the well-known CIDOC CRM ontology created to serve the specific needs of the archaeological community; and developing innovative services. The ARIADNE Catalogue is accessible via the ARIADNE Portal (http:/portal.ariadne-infrastructure.eu/) and can be queried using a powerful semantic engine based on time, space and object type facets. Datasets are kept at their original location and are managed by the owner. The Catalogue stores only dataset metadata, which are linked to the source documents that can be directly accessed once selected in the Portal.

Objectives of ARIADNEplus

The overall goal of ARIADNEplus (Niccolucci, 2018), may be summarized as: “Extending and Focusing ARIADNE”. Extending concerns the domains served and the users addressed; it has several dimensions:

- The geographic coverage, which in ARIADNE already reached almost all the European regions, by integrating in the ARIADNEplus infrastructure a greater number of archaeological partners with a particular attention to areas where the coverage was less intensive;
- The disciplinary coverage, which in ARIADNE included mainly excavation data and a few other topics as, for example, dendrochronology, by integrating in the new ARIADNEplus Data Infrastructure data produced within other archaeological subdomains, e.g. palaeoanthropology, bioarchaeology, environmental archaeology material sciences, dating and so on;
- The time-span considered, pushing back the earliest datasets included, and forwards the end-date until recent times, in practice covering the full time span of the human presence on Earth;
- The depth of database integration, exploiting the potential of well-structured datasets such as databases, and archaeological Geographic Information Systems (GIS);
- The integration of text datasets by extending the use of Text Mining through Natural Language Processing (NLP) and Named Entity Recognition (NER), previously applied only experimentally. The ARIADNEplus rule-based tool, called TEXTCROWD and tested in the EOSC framework within the EOSC-pilot project, relies on multilingual archaeological vocabularies. Its performance will be compared with a novel machine-learning implementation, also EOSC-compliant, TEXTCROWDplus. Both outperform other NLP tools as they are tailored to the archaeological framework and language.
- The research community involved. The ARIADNEplus target is to make contact with the majority of all researchers and professionals and address most of their needs;
- The service portfolio offered to users, incorporating more advanced tools for digital analysis and interpretation in the ARIADNEplus System.

[1] Author's address: Franco Niccolucci, PIN, Piazza Ciardi 25, 59100 Prato, Italy, franco.niccolucci@gmail.com

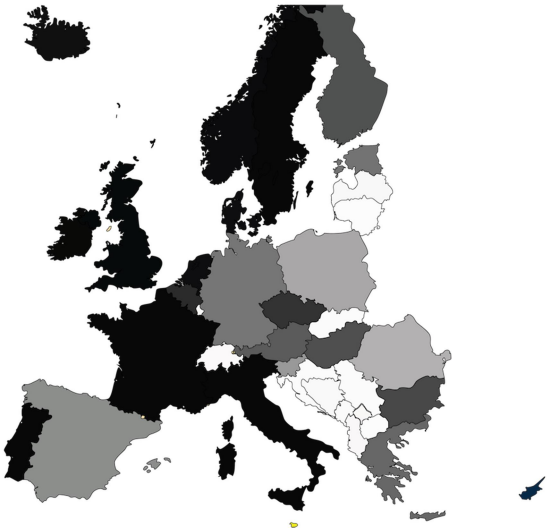

Fig. 1. Geographic coverage of ARIADNEplus. A darker colour indicates a larger number of dataset

Community impact

The completion of the ARIADNE project fostered a new approach to archaeological research. Data that before were often considered a mere support to documentation started to become the support for new investigations. Important national experiences like those of the Archaeology Data Service (ADS) in the UK had already generated the birth of similar initiatives in other countries. This stimulated the growth of a new attitude towards cooperative research, fostering the concepts of data sharing and re-use and a potential new methodology in which data were one of the pillars of archaeological investigations.

The Networking Activities planned in ARIADNEplus will support large scale adoption of this new perspective. The collaboration with associations and institutional bodies will facilitate the community penetration of the ARIADNEplus approach, as well as in the wider research community. This novel approach to archaeological research is supported by standardization, state-of-art technology and innovative services. Furthermore, with ARIADNEplus, archaeology will be able to participate in the EOSC as a protagonist.

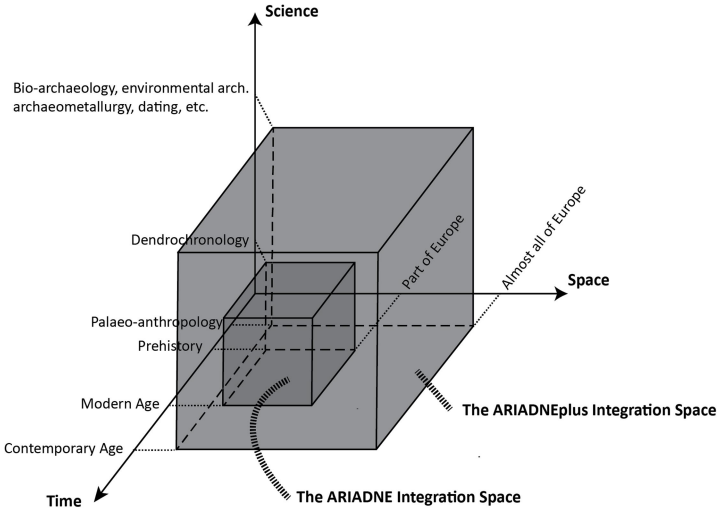

Fig. 2. How ARIADNEplus will extend its integration scope in time, space and content

References

Meghini, C., Scopigno, R., Richards, J.D., Wright, H., Geser, G., Cuy, S., Fihn, J., Fanini, B., Hollander, H., Niccolucci, F., Felicetti, A., Ronzino, P., Nurra, F., Papatheodorou, C., Gavrilis, D., Theodoridou, M. , Doerr, M., Tudhope, D. ,

Binding, C. and Vlachidis, A. (2017) “ARIADNE: A Research Infrastructure for Archaeology”. JOCCH Special Issue on Digital Infrastructure for Cultural Heritage, Part 2 Volume 10 Issue 3, August 2017, Article No. 18.

Niccolucci, F. and Richards, J.D. (2013) “ARIADNE: Advanced Research Infrastructures for Archaeological Dataset Networking in Europe”. Int. Journal of Humanities and Arts Computing, vol. 7, No. 1-2: 70-88.

Niccolucci, F. (2018) “Integrating the digital dimension into archaeological research: the ARIADNE project”. European Journal of Post-Classical Archaeology 8: 63-74.