ARIADNE infrastructure: Benefits and requirements of data providers

Guntram GESER, Salzburg Research Institute, Austria

The ARIADNE data infrastructure for archaeology in brief

ARIADNEplus (https://ariadne-infrastructure.eu) is a project funded by the European Commission to provide a digital infrastructure that enables the archaeological community to share, find and access data in dispersed repositories for use across institutional and national as well as disciplinary boundaries. The core components of the Cloud-based infrastructure are a catalogue of indexed records of various archaeological data and a data search portal, currently providing access to around 2.8 million records, and growing. The portal allows searching the data records according to the three facets of "when" (time), "where" (space), "what" (object), based on controlled vocabularies. The available data resides in the repositories of the contributors who maintain and control the data access according to their terms.

Benefits of sharing data through ARIADNE

The ARIADNE initiative is being recognised by archaeological institutions and projects as the leading integrator of archaeological datasets. Consequently, the initiative increasingly receives requests to join it as associate partner institution or project. What benefits are they looking for, and how does the sector of archaeology overall benefit from the ARIADNE data infrastructure?

These benefits can be summarised referring to the well-known FAIR data principles, which are being adopted by ever more research funders, in addition to the open data agenda, i.e. that data of publicly funded projects should also become available for other users. The ARIADNE initiative supports making archaeological data better Findable, Accessible, Interoperable, and Reusable (FAIR), not only data of one repository or database but across all who provide data to the initiative.

The key benefits of this are that ARIADNE offers archaeological institutions and projects the opportunity to describe their datasets using advanced metadata standards and, based on that, integrate the data records in the ARIADNE data search and access portal. Thereby individual datasets and their providers gain greater visibility and the potential of data discovery, access and reuse increases.

More specifically, ARIADNE enables high quality data description and integration based on the ISO standard CIDOC Conceptual Reference Model (CIDOC-CRM), implemented in the ARIADNE data

catalogue model, additional domain-specific extensions of the model (e.g., CRMarcheo for archaeological excavations), and common vocabularies, particularly the PeriodO gazetteer for spatio-temporal boundaries of cultural periods, and the Getty Art & Architecture Thesaurus (AAT) for subjects.

This standardization increases the "FAIRness" of archaeological data as the FAIR principles, for example, require that (meta)data should be based on relevant community standards (i.e., CIDOC-CRM), be described with accurate and relevant attributes (controlled vocabularies such as the AAT), and registered in a searchable resource, i.e., the ARIADNE data portal, in addition to a searchable institutional or thematic repository.

Requirements for reaping the benefits

There are of course requirements for reaping the benefits offered by ARIADNE. This short paper addresses some of the requirements, listed in Figure 1 alongside the mentioned benefits. The discussion of the requirements refers to results of two online surveys carried out by the ARIADNEplus project: The community needs survey in 2019 included questions on data sharing, access and (re)use, and received 484 useable questionnaires (ARIADNEplus, 2019). The data-related questions were a follow-up to a survey in 2013 of the original ARIADNE project with over 500 participants (ARIADNE, 2014). The second ARIADNEplus survey in 2021 focused on practices and needs of archaeological repositories regarding FAIR data, data policies, how to improve data access, and other pertinent questions. This survey received information from 60 repositories, 43 operative and 17 being set up, with the majority from around Europe (Geser et al., 2022).
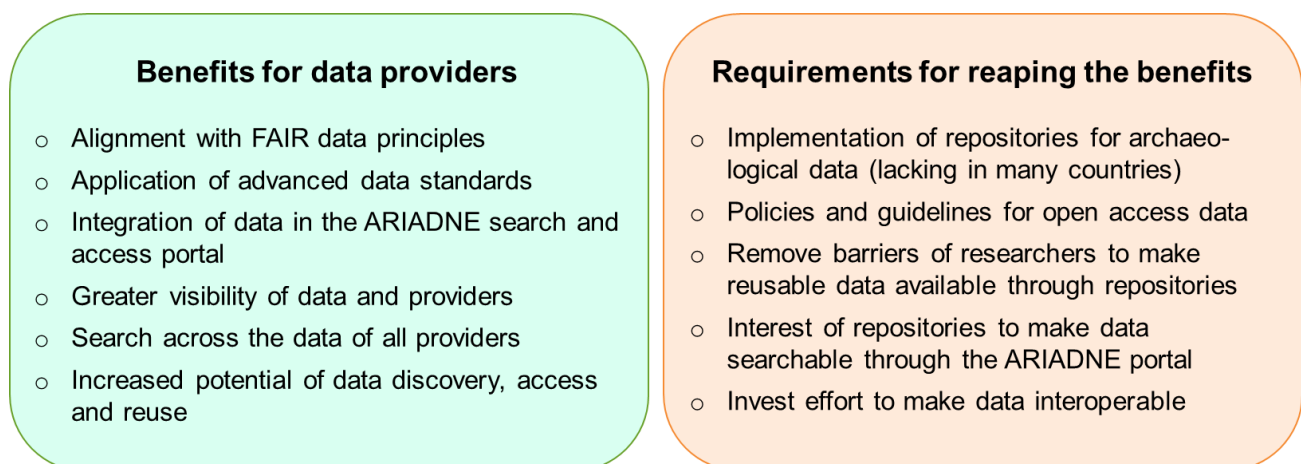
Benefits for data providers

- Alignment with FAIR data principles
- Application of advanced data standards
- Integration of data in the ARIADNE search and access portal
- Greater visibility of data and providers
- Search across the data of all providers
- Increased potential of data discovery, access and reuse

Requirements for reaping the benefits

- Implementation of repositories for archaeological data (lacking in many countries)
- Policies and guidelines for open access data
- Remove barriers of researchers to make reusable data available through repositories
- Interest of repositories to make data searchable through the ARIADNE portal
- Invest effort to make data interoperable

Fig. 1. ARIADNE infrastructure: Overview of benefits and requirements of data providers (© Guntram Geser).

Available open access repositories appropriate for archaeological data

The ARIADNE infrastructure enables discovery of, and access to, data deposited by archaeologists and heritage managers in existing repositories. However, in some European and other countries researchers and practitioners lack a repository for archaeological data that have special requirements (Wright & Richards 2018). In the ARIADNEplus community needs survey 2019, 67% of respondents perceived a lack of appropriate repositories, almost the same percentage as in the survey 2013 (66%). This issue is being addressed by ARIADNEplus' "sister project" SEADDA[1], which is a European COST Action aimed to improve the unfavourable situation.

---

[1] Saving European Archaeology from the Digital Dark Age (SEADDA), https://www.seadda.eu

In addition to support for the implementation of state-of-the-art archaeological repositories in terms of technology and data standards, there is also a need of appropriate policies and guidelines for open data access. In the ARIADNEplus survey of repositories the respondents were asked *"What would help the repository most to support open data access and reuse policies?"* The main help needed was heritage regulations to set respective policies (39 of 56 respondents) and clear guidelines by heritage authorities (36). Next came the challenge to overcome barriers of researchers to deposit open and reusable data (29), e.g., concerns about open licensing. Respondents also considered as important training of repository staff to support new policies on FAIR and open data (28).

Remove barriers of researchers to make data available through open access repositories

In the ARIADNEplus community survey 2019, around 65% of respondents said they published some to all their project data through an accessible repository, in 2013 around 50% said they did so, suggesting a significant increase by 15%. However, most data was still contained in academic publications, fieldwork reports, and supplementary material, and therefore not easily reusable.

The survey participants were also asked if they agreed with the statement: *"In the last 5 years the readiness of archaeologists to share data through publicly accessible repositories or databases increased"?* – 83.2% of respondents agreed. However, several perceived a higher awareness among archaeologists that reusable data should be made available, but little increase in readiness to do so. Others felt that the increase is taking place only slowly.

The impact of the COVID-19 pandemic on archaeology and cultural heritage (Geser, 2021) may have made archaeologists more aware of the importance of access to publicly shared data, data repositories and discovery and access services. In the repositories survey, run between June and September 2021, 24 of the 27 repositories that analysed data access reported that during the pandemic overall there was increased access, with increases ranging from 5% to over 100%.

However, research institutions and funders must do more to foster open data sharing: In the 2019 community survey, respondents considered as particularly strong barriers a lack of recognition and reward for data sharing (75.5%), intellectual property rights issues (75%), and concerns about additional work effort for providing data and metadata in the required formats (74%). These barriers must be addressed with measures such as dedicated financial support of data sharing through open access repositories and promotion of researchers and research groups that stand out in this regard.

Interest of repositories to make data searchable on the ARIADNE data portal

Repository members of ARIADNEplus and requests by others to join the project as a not funded associate partner confirm existing interest. One of the questions of the repositories survey was, *"Does the repository make metadata available to external search platforms/engines?".* Of the 60 respondents 25 said "Yes", 26 "No", and 9 "Don't know".

Many respondents (26) said that their repository does not share metadata with external search platforms, which can have different reasons, for example, a lack of metadata quality or a legacy data management system that does not support metadata harvesting. Indeed, asked *"What would help the repository most for improving data access?"*, most important of seven suggested options across the 60 respondents were to improve the quality of metadata (34), improve or replace the existing data management system (30), and provide metadata to external search platforms (27).

ARIADNE provides a solution for the metadata quality issue by supporting data description using advanced domain standards and, based on that, integration in the ARIADNE data search and access portal. This of course requires some work by the repositories, which includes mapping their data-bases to the CIDOC-CRM based ARIADNE data catalogue model (or extended versions such as CRMarcheo for archaeological excavations), and aligning terms of a local thesaurus or term list to concepts of the large and multilingual Getty Art & Architecture Thesaurus. For this work, ARIAD-NEplus partners provide the required tools and support in using them. Based on this and some related work data records of the providers can be aggregated and integrated in the ARIADNE portal for data search and access across the different providers.

Brief summary

The ARIADNE infrastructure provides search and access to data from research on archaeological and cultural heritage deposited in repositories. It enables the archaeological community to share, find and access data in dispersed repositories for use across institutional and national as well as disciplinary boundaries. Integration of (meta)data in the ARIADNE search and access portal is based on FAIR data principles and application of advanced domain standards. The integration enables greater visibility of available datasets and their providers and increases the potential of data discovery, access and reuse.

Among the requirements to reap the benefits offered by the ARIADNE initiative are available repositories that are appropriate for archaeological data, open data policies and guidelines, especially of heritage authorities, and removal of barriers of researchers to make reusable data available through open access repositories. Particularly required is the interest of repositories to make interoperable data searchable through the ARIADNE portal, i.e., implement the data standards the ARIADNE initiative promotes to increase the "FAIRness" of archaeological data.

Funding

References

ARIADNE (2014). First Report on Users' Needs, D2.1, April 2014. Available at http://legacy.ariadne-infrastructure.eu/resources-2/deliverables/ (Accessed: 28 June 2022).

ARIADNEplus (2019). Initial Report on Community Needs, D2.1, October 2019. Available at https://ariadne-infrastructure.eu/wp-content/uploads/2019/11/ARIADNEplus_D2.1_Initial-Report-on-Community-Needs-1.pdf (Accessed: 28 June 2022).

Geser, G. (2021). Impact of COVID-19 on Archaeology and Cultural Heritage, ARIADNEplus, 29 October 2021. Available at https://ariadne-infrastructure.eu/wp-content/uploads/2021/11/COVID-19_impact-archaeology-and-cultural-heritage_29Oct2021.pdf (Accessed: 28 June 2022).

Geser, G., Richards, J.D., Massara, F. and Wright, H. (2022). 'Data Management Policies and Practices of Digital Archaeological Repositories', Internet Archaeology 59. DOI: 10.11141/ia.59.2

Wright, H. and Richards, J.D. (2018). 'Reflections on Collaborative Archaeology and Large-scale Online Research Infrastructures', Journal of Field Archaeology, 43, supp1, S60–S67. DOI: 10.1080/00934690.2018.1511960