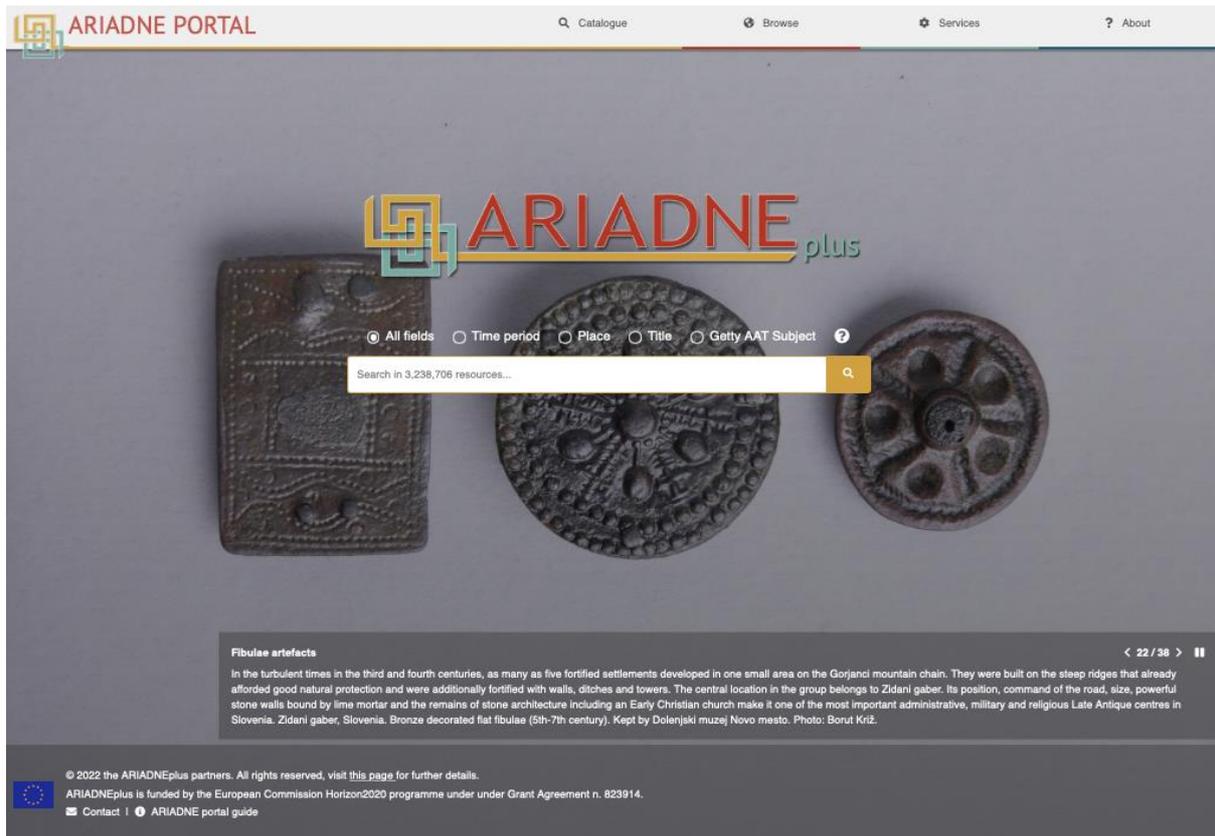


## Joined Up Thinking: The challenges of large-scale data integration

Prof Julian D. RICHARDS, Director, Archaeology Data Service, University of York

Research e-infrastructures, digital archives, and data services have become important pillars of scientific enterprise that in recent decades has become ever more collaborative, distributed, and data intensive. The archaeological research community has been an early adopter of digital tools for data acquisition, organization, analysis, and presentation of research results of individual projects. However, the provision of e-infrastructure and services for data sharing, discovery, access, and re-use has lagged behind. This situation has been addressed by ARIADNE. This EU-funded network has developed an e-infrastructure that enables data providers to register and provide access to their digital resources through the ARIADNE data portal, facilitating discovery, access, and research.

ARIADNE has aggregated resources from over 30 data providers, spanning over 40 countries and 4 continents. The ARIADNE portal provides online access to over 3.2 million research resources, with the number increasing daily. In the current phase of the project, ARIADNEplus, there has been an explicit effort to increase the temporal and geographical range of resources, and to extend the thematic coverage. However, it is equally important that we do not simply create another data silo, just larger than previous efforts. The ARIADNE Knowledge Base is therefore based upon Linked Open Data technologies and is underpinned by a flexible and extensible architecture. We have also been keen not to “make a great heap” of all the data and, learning from previous data aggregation projects, we have defined a sub-set of the CIDOC-CRM as a strict ontology and paid close attention to data standards and controlled vocabularies to achieve a high level of interoperability. This paper will discuss some of the challenges of large-scale data integration and describe the approaches adopted to ensure that the ARIADNE Knowledge Base is an effective tool for resource discovery. It will seek to demonstrate how the portal can also be used to address major research questions, and how the Linked Open Data approach provides flexibility, with the portal architecture enabling multiple combinations and presentations of the same underpinning data.



## References

- Meghini, C., Scopigno, R., Richards, J.D. et al. 2017. "ARIADNE: A Research Infrastructure for Archaeology", *Journal of Computing and Cultural Heritage* 10.3, 1-27, <https://doi.org/10.1145/3064527>
- Richards, J.D. and F. Niccolucci (eds.) 2019. *The ARIADNE Impact*. Archaeolingua, <https://doi.org/10.5281/zenodo.4319058>
- Wright, H. and Richards, J.D. 2018. "Reflections on Collaborative Archaeology and Large-Scale Online Research Infrastructures", *Journal of Field Archaeology*, 43:sup1, S60-S67, <https://doi.org/10.1080/00934690.2018.1511960>