



Intro to Saving European Archaeology from the Digital Dark Age and the FAIR Principles

National Workshop – Bucharest, 18 May 2023

Holly Wright | Archaeology Data Service | University of York



ARCHAEOLOGY
DATA SERVICE

UNIVERSITY *of* York



COST Action SEADDA (CA18128)
is funded by the Horizon 2020
Framework Programme of the
European Union.



The Archaeology Data Service

- Set up in 1996
- Based at the University of York



Mission

Supporting research, learning and teaching with free, high quality and dependable digital resources

- Preserve data
- Free online access to data
- Guidance and support for data creators
- Research

<http://guides.archaeologydataservice.ac.uk>

Lessons from ARIADNE

- EU FP7 project under the ‘Infrastructures’ theme, including 23 European partners across 16 countries
- Four-year project: ended January 2017
- 17 Workpackages – primary deliverable: collection level metadata aggregation portal
- Coordinator: PIN Scri - Polo Universitario "Città di Prato" at the University of Florence
- Deputy Coordinator: ADS



ARIADNE

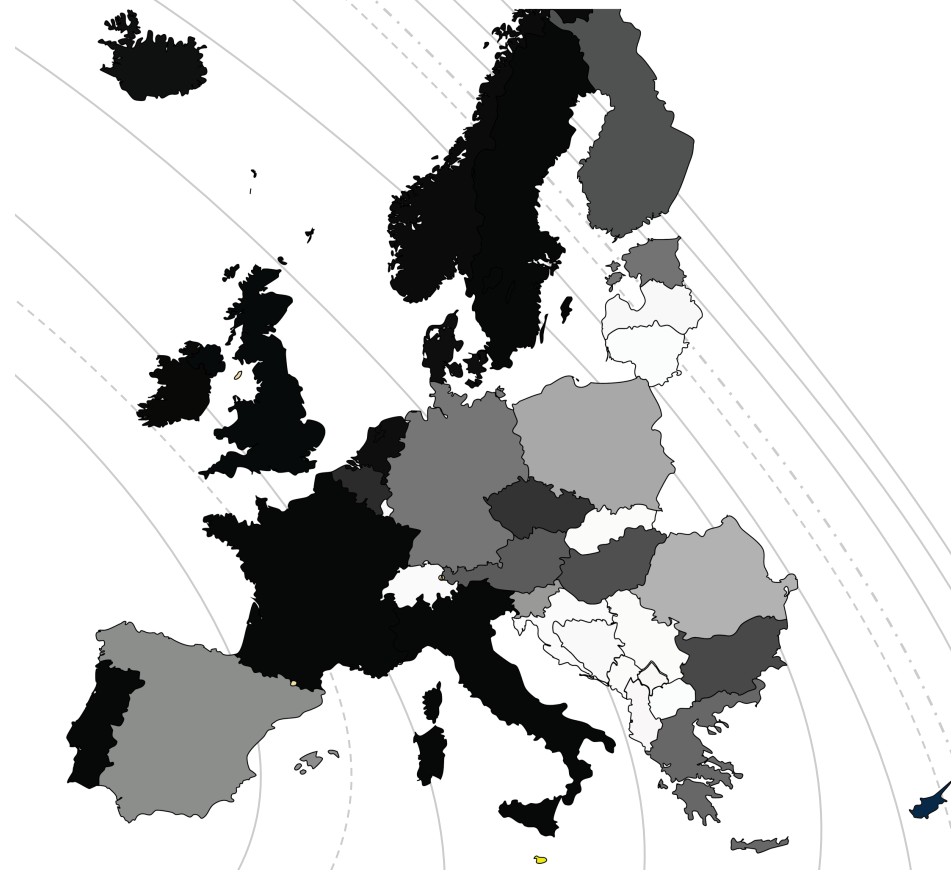
- 9 ICT organisations
- 14 archaeological organisations
- 15 Associate partners
- External community building
 - Transnational access
 - Training events
 - Special interest groups

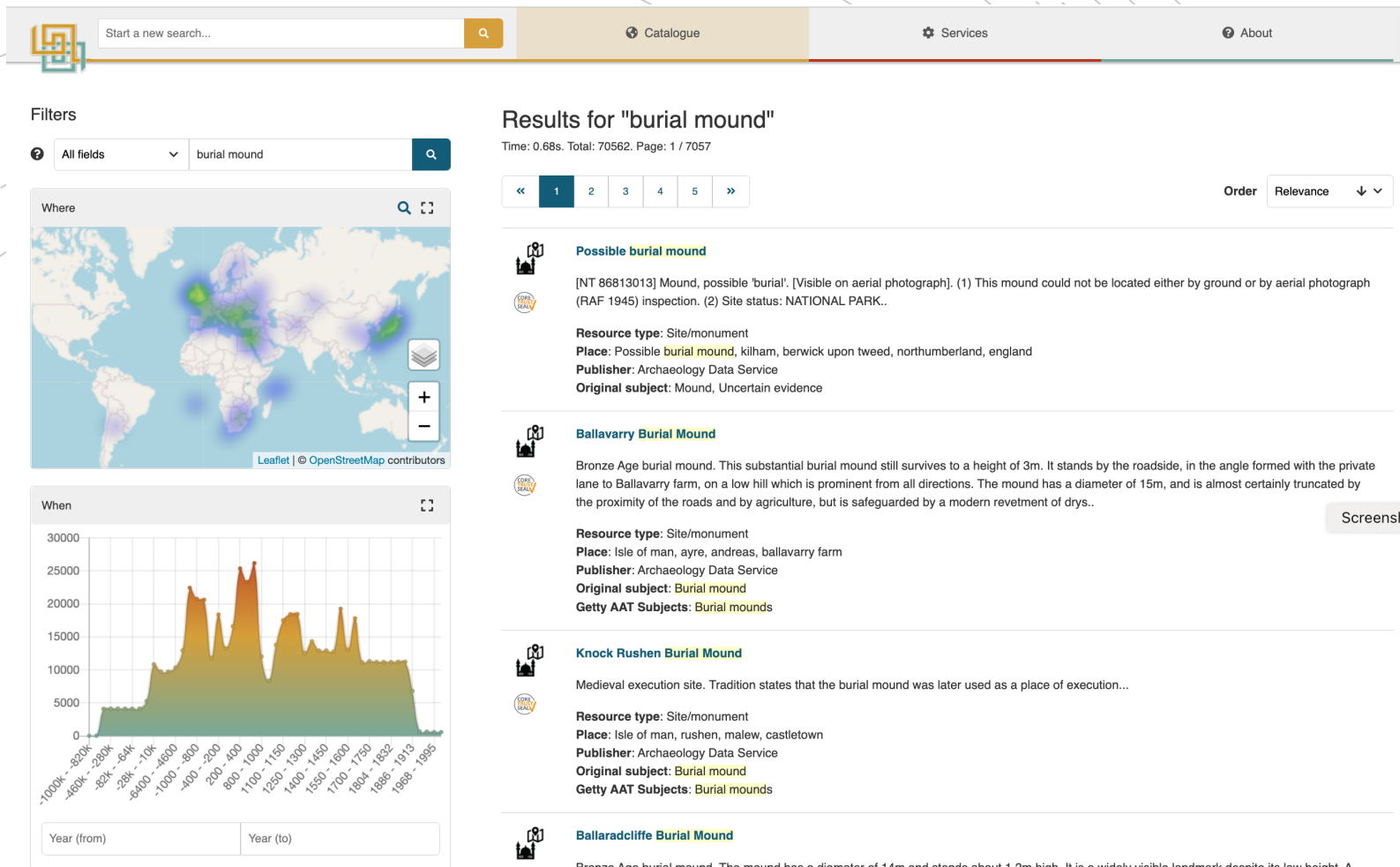


ARIADNEplus

- 41 partners
- 28 countries
- Majority are archaeological partners
- Expanding participation
- Extending thematic coverage

<https://ariadne-infrastructure.eu/>





The screenshot shows the SEADDA portal interface. At the top, there is a search bar with the text "Start a new search..." and a search icon. To the right of the search bar are navigation links for "Catalogue", "Services", and "About".

Below the search bar, there is a "Filters" section. It includes a dropdown menu set to "All fields" and a search box containing the text "burial mound". Below this is a "Where" section with a world map showing search results as colored dots. Below the map is a "When" section with a bar chart showing the distribution of results over time. The x-axis of the chart is labeled "Year (from)" and "Year (to)" and ranges from -1000k to 1995. The y-axis represents the number of results, ranging from 0 to 30,000. The chart shows a significant increase in results starting around 1000 BC, peaking around 1000 AD, and then declining.

The main content area displays "Results for 'burial mound'" with a total of 70562 results. The results are listed in a table format, with each entry including a title, a description, and metadata. The first three results are:

- Possible burial mound**: [NT 86813013] Mound, possible 'burial'. [Visible on aerial photograph]. (1) This mound could not be located either by ground or by aerial photograph (RAF 1945) inspection. (2) Site status: NATIONAL PARK..
Resource type: Site/monument
Place: Possible burial mound, kilham, berwick upon tweed, northumberland, england
Publisher: Archaeology Data Service
Original subject: Mound, Uncertain evidence
- Ballavarry Burial Mound**: Bronze Age burial mound. This substantial burial mound still survives to a height of 3m. It stands by the roadside, in the angle formed with the private lane to Ballavarry farm, on a low hill which is prominent from all directions. The mound has a diameter of 15m, and is almost certainly truncated by the proximity of the roads and by agriculture, but is safeguarded by a modern revetment of drys..
Resource type: Site/monument
Place: Isle of man, ayre, andreas, ballavarry farm
Publisher: Archaeology Data Service
Original subject: Burial mound
Getty AAT Subjects: Burial mounds
- Knock Rushen Burial Mound**: Medieval execution site. Tradition states that the burial mound was later used as a place of execution...
Resource type: Site/monument
Place: Isle of man, rushen, malew, castletown
Publisher: Archaeology Data Service
Original subject: Burial mound
Getty AAT Subjects: Burial mounds

The fourth result is partially visible: **Ballaradcliffe Burial Mound**. Below the results, there is a "Screenshots" button.

Internal Capacity Building

- Paper was given in 2015 by partners in Slovenia and Ireland talking about “have’s and have nots” when it comes to having an acceptable place for their archaeological data in the long term
- Plans were made to offer data management workshops to partners in their own countries - Austria and Slovenia both accepted

Data management workshops

Local ARIADNE partners worked hard to make sure governmental and institutional stakeholders attended, and took on understanding and presenting the state of play in their country



Data management workshops

Level of interest was much higher than expected:

- Pressure from funders to deposit data into an open repository
- Pressure from institutions to deposit data into an open repository
- Funding available; but no guidance
- No appropriate options for archaeological data
- Conversations continued about how to take expertise within ARIADNE and collaborate beyond current network
- Working together was going to be critical to moving forward

Challenges for Archaeology

- Archaeological data often derived from non-repeatable interventions
- Digital data more fragile and subject to obsolescence
- Risk losing a generation of research – Urgency!



COST Action SEADDA :

Saving European Archaeology from the Digital Dark Age

Funding (four years) for networking: meetings, training sessions, scientific missions, open access publications, with members representing 35 countries



- Coordinate information collection to understand the current state-of-the-art regarding the preservation, dissemination and re-use of archaeological data.
- Develop a common understanding of international best practice for preservation, dissemination and re-use of archaeological data
- Foster knowledge exchange around international best practice for preservation, dissemination and re-use of archaeological data

Stewardship of archaeological data

Objective: To bring together members with varying levels of experience to share their successes and challenges around the stewardship of archaeological data to create a sub-network.

Practical and ethical considerations will be explored:

- Encouragements and resistances to sharing data and making it openly accessible within archaeology
- Who should be responsible in short and long-term for preservation

The starting point for those who wish to begin or progress dialogue in their region or country.

National workshops!

Planning for Archiving

Objective: To identify the practical and technical issues surrounding the creation of an appropriate repository for archaeological data

- Understanding hardware and software options
- Management structures
- Training of archivists

The WG identifies existing best practice, changing future needs, and pragmatic technical and structural solutions.

Preservation and Dissemination Best Practice

Objective: Understand current international best practice for archiving and dissemination, and implementation by existing repositories.

- Open Archival Information System (OAIS) model
- The FAIR Principles
- Repository accreditation
- Cost modelling
- Dealing with data types

Will bring together archaeological digital archivists to share current practice, and survey future trends to understand the changing archaeological and digital landscapes (domain and technology watch).

Use and Re-Use of Archaeological Data

Objective: To understand how to optimise archives and interfaces to maximise the use and re-use of archaeological data. To explore how archaeological archives can better respond to user needs, and ways to document and understand both quantitative and qualitative re-use.

- Explore barriers to re-use, such as IPR and licencing
- Explore design of underlying data structures and their interfaces

Will focus on initiatives like the FAIR Principles and technologies that improve and optimise searching, issues around how data is created, organised and disseminated, different options for interface design, and developing best practice around qualitative re-use.

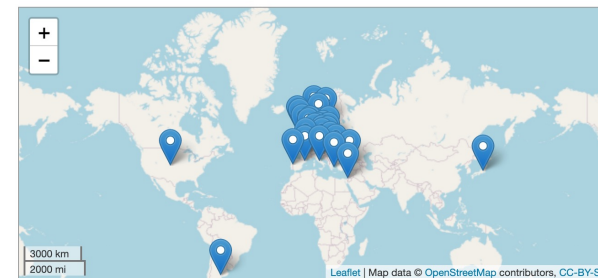
Open Access Publications

- Internet Archaeology SEADDA themed issue *Digital Archiving in Archaeology: The state of the art*
- 28 papers representing the current state of archiving in 28 nations and states.
- Also just published: *Data Management Policies and Practices of Digital Archaeological Repositories*


Digital Archiving in Archaeology: The State of the Art. Introduction

Julian D. Richards, Ulf Jakobsson, David Novák, Benjamin Štular and Holly Wright


Cite this as: Richards, J.D., Jakobsson, U., Novák, D., Štular, B. and Wright, H. 2021 Digital Archiving in Archaeology: The State of the Art. Introduction, *Internet Archaeology* 58. <https://doi.org/10.11141/ia.58.23>




The advent of ubiquitous computing has created a golden age for archaeological researchers and participating publics, but the price is a digital resource that is now in jeopardy. The archaeological record, in digital form, is at risk not simply from obsolescence and media failure, but the domain is also unable to fully participate in Open Data. Without swift and informed

Corresponding author: Julian D. Richards 
julian.richards@york.ac.uk
Archaeology Data Service

Ulf Jakobsson
Swedish National Data Service

David Novák 
Institute of Archaeology, Czech Academy of Sciences

Benjamin Štular 
ZRC-SAZU, Ljubljana, Slovenia

Holly Wright 
Archaeology Data Service

<https://intarch.ac.uk/journal/issue59/2/index.html>



<https://www.seadda.eu/>

What are the FAIR Principles?

In 2016, the ‘FAIR Guiding Principles for scientific data management and stewardship’ were published in Scientific Data. The authors intended to provide guidelines to improve the **F**indability, **A**ccessibility, **I**nteroperability, and **R**euse of digital assets. ***The principles emphasise machine-actionability (i.e., the capacity of computational systems to find, access, interoperate, and reuse data with little or no human intervention)*** because humans increasingly rely on computational support to deal with data as a result of the increase in volume, complexity, and creation speed of data.

The first step in (re)using data is to find them. Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services.

- F1.** (Meta)data are assigned a globally unique and persistent identifier
- F2.** Data are described with rich metadata
- F3.** Metadata clearly and explicitly include the identifier of the data they describe
- F4.** (Meta)data are registered or indexed in a searchable resource

Once the user finds the required data, she/he needs to know how can they be accessed, possibly including authentication and authorisation.

A1. (Meta)data are retrievable by their identifier using a standardised communications protocol

A1.1 The protocol is open, free, and universally implementable

A1.2 The protocol allows for an authentication and authorisation procedure where necessary

A2. Metadata are accessible, even when the data are no longer available

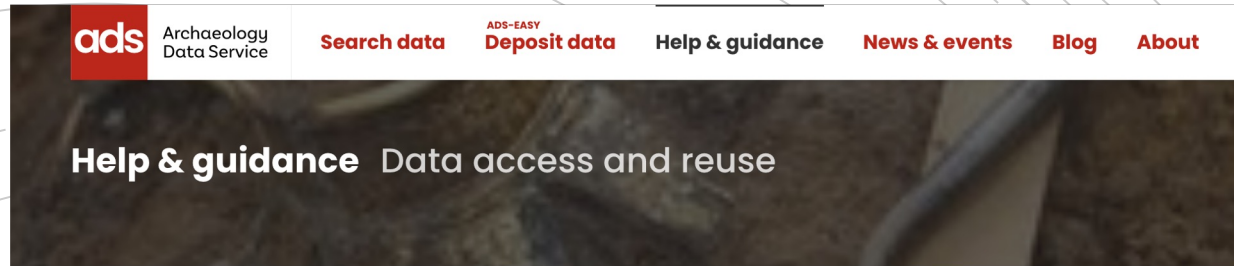
The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing.

- I1.** (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2.** (Meta)data use vocabularies that follow FAIR principles
- I3.** (Meta)data include qualified references to other (meta)data

The ultimate goal of FAIR is to optimise the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings.

- R1.** (Meta)data are richly described with a plurality of accurate and relevant attributes
 - R1.1.** (Meta)data are released with a clear and accessible data usage license
 - R1.2.** (Meta)data are associated with detailed provenance
 - R1.3.** (Meta)data meet domain-relevant community standards

What does this look like in practice?



Data access and reuse

■ FAIR data

Identifying copyright

Data reuse case studies

Digital Object Identifiers (DOI)

FAIR data

The ADS is an advocate for FAIR and the FAIR principles for data stewardship. As such the ADS recognise that while preservation and dissemination of data remain of core importance, stewardship should also include demonstrable quantitative and qualitative evidence for data reuse. The ADS is actively investigating how the datasets it curates can be fully compliant with the FAIR principles and is working within [SSHOC](#), [ARIADNEplus](#) and [E-RIHS](#) to promote this.

As a result when you deposit your datasets with the ADS, you can be confident that your data becomes FAIR data.

What is FAIR Data?

The [FAIR Principles](#) provide an important framework to evaluate and publish data in order to facilitate discovery, provide sustainable access to resources, and encourage and enable better sharing and reuse of data. To achieve these goals the core principles emphasise:

<https://archaeologydataservice.ac.uk/help-guidance/data-reuse/fair-data/>

Interoperable

11. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

- All resource discovery metadata is made available using a qualified Dublin Core in RDF/XML through the [ADS Linked Data repository](#).
- [External services](#) also consume and disseminate metadata.

12. (Meta)data use vocabularies that follow FAIR principles.

For a wider discussion on the vocabularies used in ADS metadata see our [Strategy and Standards](#) page.

The ADS uses a variety of sustainable, open vocabularies to qualitatively classify and identify resources and datasets, including:

- [Heritage Data](#) vocabularies, including those provided by the Forum on Information Standards in Heritage (FISH), Historic England (HE), Historic Environment Scotland (HES), and the Royal Commission on Ancient & Historical Monuments of Wales (RCAHMW)
- [Library of Congress Subject Headings](#) (LCSH)
- [Marine Environmental Data and Information Network](#) (MEDIN)
- [Getty Thesaurus of Geographic Names](#) (TGN)

The ADS also utilises recognised technical vocabularies to denote and categorise preservation activities.

- [PREservation Metadata: Implementation Strategies \(PREMIS\)](#)
- Getty metadata types ([Baca 2016](#))

13. (Meta)data include qualified references to other (meta)data.

- The ADS supports the qualified referencing with and between publications, datasets and resources. Where available the repository uses sustainable referencing, e.g. DOIs.

- How to make data **Findable, Accessible** and **Interoperable** are well understood, with examples of well-implemented methodologies and technologies
- Still lots of work to do on **Reusable**: Can measure quantitative reuse with web stats, but how to measure qualitative reuse is the next frontier
- **FAIR** makes each element of equal importance
- **FAIR** Principles are just a useful lens for understanding your own situation with regard to current best practice

Holly Wright

holly.wright@york.ac.uk

Archaeology Data Service

<http://archaeologydataservice.ac.uk>

SEADDA COST Action

<http://seadda.eu>



@seadda_cost