Outline

• Research data and digital corpora
  • How should we be developing them?
  • The Bigger Picture – infrastructure frameworks

• Importance of FAIR data

• Research e-infrastructures, and ADS role:
  • E-RIHS (European Research Infrastructure for Heritage Science)
  • SSHOC (Social Sciences and Humanities Open Cloud)
  • ARIADNEplus
    • Archaeologists need to engage with these initiatives
    • Introducing the ARIADNEplus taskforces
Introduction to FAIR principles

Findable  Easy to find by both humans and computer systems;

Accessible  Stored for long term such that they can be easily accessed and/or downloaded with well-defined license and access conditions;

Interoperable  Ready to be combined with other datasets by humans as well as computer systems;

Reusable  Ready to be used for future research.
• Being FAIR when archaeological information is MEAN:
  • Miscellaneous
  • Exceptional
  • Arbitrary
  • Nonconformist

(Isto Huvila, Centre for Digital Heritage conference, University of Leiden 2017)
Guidelines to *fairify* your data management
PARTHENOS FAIR guidelines

20 GUIDELINES to FAIRify data management and make data reusable

Invest in people and infrastructure

An important prerequisite to be able to implement the rest of the nineteen guidelines in this guide, is to invest in data infrastructures and in hiring and educating data experts.

Get acquainted with best practices in research data management. Check out the PARTHENOS training modules on data management or have a look at the CESSNA Data Management Expert Guide.

Invest in hiring and educating data experts and define a budget for making investments in technical infrastructure and staff.

1. Invest in people and infrastructure

2. Find the right people

Locate and hire data stewards and inform team of your responsibilities. Provide them with training and ongoing support. Make sure they have access to technical resources.

3. Use persistent identifiers

Link your data to persistent identifiers in any other data. Use a unique identifier for each piece of data. Make sure that these identifiers are openly accessible. Provide an identifier for each piece of data and make sure that it is persistent.

4. Use persistent author identifiers

Use a persistent author identifier (e.g., ORCID, ORCID) to increase the visibility and accessibility of your data. Make sure that these identifiers are openly accessible. Provide a persistent identifier for each presenter.

5. Choose an appropriate metadata schema

Choose a metadata schema that is appropriate for the data that you need to publish. Make sure that the schema is openly accessible. Provide a persistent identifier for each metadata schema.

6. Cite research data

If research data have a persistent identifier, use it in your research. Provide clear instructions on how to cite research data.

7. Choose a trustworthy repository

A trustworthy repository offers a trustworthy home for data. Make sure that the repository is open and persistent. Make sure that the repository is well-documented and that the data is citable.

8. Make your data accessible through a trustworthy repository

Publish your data through a trustworthy repository. Make sure that the repository is well-documented and that the data is citable.

9. Choose a clear and accessible file format

Choose a file format that is clear and accessible to a wide range of software and applications.

10. Make sure that your data is accessible

Make sure that your data is accessible through a trustworthy repository. Make sure that the repository is well-documented and that the data is citable.

11. Choose a clear and accessible file format

Choose a file format that is clear and accessible to a wide range of software and applications.

12. Use standardized exchange protocols

By using standardized exchange protocols, research infrastructures can share, transfer, and access data efficiently and reliably. Make sure that the protocols are open and persistent. Provide a persistent identifier for each exchange protocol.

13. Use standardized data exchange protocols

Use standardized data exchange protocols to ensure that data can be shared and reused across different platforms and applications.

14. Use standardized data exchange protocols

Use standardized data exchange protocols to ensure that data can be shared and reused across different platforms and applications.

15. Choose a clear and accessible file format

Choose a file format that is clear and accessible to a wide range of software and applications.

16. Make sure that your data is accessible

Make sure that your data is accessible through a trustworthy repository. Make sure that the repository is well-documented and that the data is citable.

17. Choose a clear and accessible file format

Choose a file format that is clear and accessible to a wide range of software and applications.

18. Use standardized exchange protocols

By using standardized exchange protocols, research infrastructures can share, transfer, and access data efficiently and reliably. Make sure that the protocols are open and persistent. Provide a persistent identifier for each exchange protocol.

19. Use standardized data exchange protocols

Use standardized data exchange protocols to ensure that data can be shared and reused across different platforms and applications.

20. Choose a clear and accessible file format

Choose a file format that is clear and accessible to a wide range of software and applications.

21. Use persistent author identifiers

Use a persistent author identifier (e.g., ORCID) to increase the visibility and accessibility of your data. Make sure that these identifiers are openly accessible. Provide a persistent identifier for each presenter.

22. Reference author identifiers in the metadata

Reference author identifiers in the metadata. Make sure that these identifiers are openly accessible. Provide a persistent identifier for each recommender.

23. Use persistent author identifiers

Use a persistent author identifier (e.g., ORCID) to increase the visibility and accessibility of your data. Make sure that these identifiers are openly accessible. Provide a persistent identifier for each presenter.

24. Reference author identifiers in the metadata

Reference author identifiers in the metadata. Make sure that these identifiers are openly accessible. Provide a persistent identifier for each recommender.

25. Use persistent author identifiers

Use a persistent author identifier (e.g., ORCID) to increase the visibility and accessibility of your data. Make sure that these identifiers are openly accessible. Provide a persistent identifier for each presenter.

26. Reference author identifiers in the metadata

Reference author identifiers in the metadata. Make sure that these identifiers are openly accessible. Provide a persistent identifier for each recommender.

27. Use persistent author identifiers

Use a persistent author identifier (e.g., ORCID) to increase the visibility and accessibility of your data. Make sure that these identifiers are openly accessible. Provide a persistent identifier for each presenter.

28. Reference author identifiers in the metadata

Reference author identifiers in the metadata. Make sure that these identifiers are openly accessible. Provide a persistent identifier for each recommender.

29. Use persistent author identifiers

Use a persistent author identifier (e.g., ORCID) to increase the visibility and accessibility of your data. Make sure that these identifiers are openly accessible. Provide a persistent identifier for each presenter.

30. Reference author identifiers in the metadata

Reference author identifiers in the metadata. Make sure that these identifiers are openly accessible. Provide a persistent identifier for each recommender.
Research e-infrastructures

E-RIHS
EUROPEAN RESEARCH INFRASTRUCTURE FOR HERITAGE SCIENCE

SSHOC
social sciences & humanities open cloud

EOSC-hub

ARIADNE plus

http://archaeologydataservice.ac.uk
E-RIHS
EUROPEAN RESEARCH INFRASTRUCTURE
FOR HERITAGE SCIENCE

synchrotron, laser, ion beam, neutron, microscopy
dating, biogeochemistry
genomics, proteomics

FIXLAB
Large-scale facilities

DIGILAB
Data Infrastructure
data documentation and sharing
high performance computing

MOLAB
portable instruments
material sampling

Fleet of advanced mobile instruments

ARCHLAB
Scientific archives
access to scientific archives
access to reference collections

http://archaeologydataservice.ac.uk
DIGILAB
Digital data and tools: Virtual access to scientific data concerning tangible heritage
• **(Task 3.3) Financial aspects of data policy and management**
  Heritage data sets are potentially very large and are collected in a variety of formats, including born-digital. However, they have a long term re-use value and European funding bodies are increasingly adopting Open Data policies. This can lead to significant long term costs – an aspect that is rarely considered. This task explores the cost models for the long term preservation of heritage data.

• **(Task 5.4) Data curation**
  Issues concerning data curation for heritage science:
  • data quality assurance
  • data life-cycle
  • data management and preservation
  The task defines policies to be adopted and will provide guidelines for researchers, e.g. for the creation of data plans within research projects.
Social Sciences and Humanities Open Cloud

Realising the Social Sciences and Humanities part of the European Open Science Cloud

http://archaeologydataservice.ac.uk
SSHOC Partners

Representing the archaeological community and ARIADNEplus
SSHOC consortium covers the whole data cycle

© Following the FAIR Principles

http://archaeologydataservice.ac.uk
SSHOC and EOSC

Elements of SSHOC and how work comes together

- Research (data) communities
- Training
- Governance

- e-Infrastructures
- Innovation
- Tools
- Marketplace

This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782

http://archaeologydataservice.ac.uk
Task 5.6 Issues in providing Open Data in Heritage Science and Archaeology

Complex Heritage Science datasets present special accessibility and interoperability issues.

Archaeological and heritage data may provide the only archival record of heritage which is destroyed in the course of the research. In different European countries there are varying approaches to open access for archaeological data reflecting different legal protection systems, sensitivities surrounding site location, and different attitudes to data collection e.g. metal-detecting by members of the public.

This task will enable progress in data sharing in this domain and will be applicable to other domains handling diverse interdisciplinary datasets.

http://archaeologydataservice.ac.uk
Deputy Coordinator
Lead archaeological partner
Extending geographically

- **ARIADNE:**
  - 24 partners; 18 countries

- **ARIADNEplus:**
  - 41 partners; 27 countries

Plus: Argentina, Japan & USA
Extending thematically
## ARIADNEplus special interest groups

<table>
<thead>
<tr>
<th>Field</th>
<th>Institution/Location</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Paleo-anthropology</strong> (CENIEH, Spain)</td>
<td><strong>Remote Sensing</strong> (ZRC-SAZU, Slovenia)</td>
</tr>
<tr>
<td><strong>Bio-archaeology and Ancient DNA</strong> (FORTH, Crete)</td>
<td><strong>Standing Structures</strong> (LNEC, Portugal)</td>
</tr>
<tr>
<td><strong>Archaeological finds made by general public</strong> (Aarhus University, Denmark)</td>
<td><strong>Spatio-temporal data</strong> (ARUP-CAS, Czech Republic)</td>
</tr>
<tr>
<td><strong>Environmental Archaeology</strong> (Umea University, Sweden : SEAD)</td>
<td><strong>Maritime and underwater archaeology</strong> (DGPC, Portugal)</td>
</tr>
<tr>
<td><strong>Inorganic Materials Study</strong> (INFN, Italy)</td>
<td><strong>Archaeological fieldwork</strong> (INRAP, France)</td>
</tr>
<tr>
<td><strong>Field Survey</strong> (University of Groningen, Netherlands)</td>
<td><strong>Inscriptions</strong> (University of Barcelona, Spain)</td>
</tr>
<tr>
<td><strong>Burials</strong> (OAEW, Austria)</td>
<td><strong>Dating</strong> (INFN, Italy)</td>
</tr>
</tbody>
</table>
ARIADNEplus and FAIR data

- Good practice in archaeological data management
- ARIADNEplus policy support tools
  - A DMP flexible template & domain protocol
  - A policy wizard explaining the main principles concerning archaeological data management
  - A standardization wizard, documenting major standards in the archaeological domain as well as authority files such as thesauri, reference collections, gazetteers etc
- Providing guidelines and support on repository creation and management
- Providing guidelines and support on repository quality control
- Managing FAIRness of archaeological data and IPR
- Training on FAIR Data Management

http://archaeologydataservice.ac.uk
Thank you for listening

E-RIHS PP is a Horizon 2020 project funded by the European Commission under Grant Agreement n. 739503

SSHOC is a Horizon 2020 project funded by the European Commission under Grant Agreement n. 823782

ARIADNEplus is a Horizon 2020 project funded by the European Commission under Grant Agreement n. 823914.

The views and opinions expressed in this presentation are the sole responsibility of the author and do not necessarily reflect the views of the European Commission.

Contact: julian.richards@york.ac.uk
Twitter: @julian62523002

http://archaeologydataservice.ac.uk