



D2.4: Final Innovation Agenda and Action Plan

Author:

Guntram Geser (SRFG)
Franco Niccolucci (PIN)



Ariadne is funded by the European Commission's
7th Framework Programme.

Version: 1.0

November 2016

About this document

This document is a contractual deliverable of the ARIADNE - Advanced Research Infrastructure for Archaeological Dataset Networking in Europe project. The report presents the Final Innovation Agenda and Action Plan (D2.4) that has been produced under WP2 “Community Building and Innovation”. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7-INFRASTRUCTURES-2012-1) under grant agreement n° 313193. But the content of this document cannot be considered to reflect the views of the European Commission.

Partner in charge of the deliverable:	Salzburg Research Forschungsgesellschaft mbH (SRFG)
Authors:	Guntram Geser (SRFG) Franco Niccolucci (PIN)
Contributors:	Contributions by researchers of all project partners
Quality review:	Holly Wright (UoY-ADS)

Table of contents

1	Introduction and overview	5
1.1	Objectives, stakeholders and ARIADNE’s role	5
1.2	The preliminary report as background	7
1.3	Promoting innovation over the next 5-10 years	8
1.4	ARIADNE sustainability planning	9
2	Overview of the innovation agenda and suggested actions	11
2.1	General, cross-cutting requirements	11
2.2	5-year innovation agenda and suggested actions	12
2.3	10-year innovation agenda and suggested actions	16
3	Detailed innovation agenda and action plan	19
3.1	General requirements in the 5-10 year horizons	19
3.1.1	Capacity building in open sharing of research data	20
3.1.2	Coordination across the different levels of e-infrastructure	20
3.1.3	Overview of innovation actors and activities	21
3.2	5-year innovation horizon	22
3.2.1	Focus area 1 – Research e-infrastructures and digital resources	23
3.2.2	Focus area 2 – Culture of open sharing and re-use of data	30
3.2.3	Focus area 3 – Data archives and curation of archaeological research data	34
3.2.4	Focus area 4 – Capacity building for open data sharing	38
3.2.5	Focus area 5 – Providing services and enabling novel applications	41
3.3	10-year innovation horizon	45
3.3.1	Towards open and innovative e-research	45
3.3.2	Shift from data- to research-focused innovations	46
3.3.3	Innovation topics and suggested actions	47
4	ARIADNE sustainability planning	53
4.1	Introduction and overview	53
4.2	FAIR data and e-infrastructures	53
4.2.1	Open access to research results	53
4.2.2	E-infrastructures and the open data challenge	55
4.3	The ARIADNE Integrating Activity for archaeology	56
4.3.1	ARIADNE project activities	56
4.3.2	ARIADNE data-related achievements	56
4.4	Sustainability of individual activities	57
4.4.1	Research activities	57
4.4.2	Networking and community building	58
4.4.3	Transnational access – online	58
4.4.4	Transnational access – training	59

4.5	Maintenance and extension of ARIADNE services	60
4.5.1	Variants of maintenance and extension	60
4.5.2	Basic sustainability solution	61
4.6	Sustainability based on medium to long-term extension scenarios	62
4.7	Implementing sustainability	66
5	References	68

1 Introduction and overview

The introduction to the Final Innovation Agenda and Action Plan briefly addresses the goals of ARIADNE, the objectives of the agenda and action plan, and the stakeholders and beneficiaries of the proposed activities. Also ARIADNE's roles in the activities are addressed. These can be summarised as helping others to make a difference with regard to progress and innovation in archaeological research based on better access to and (re-)usability of research data. Furthermore, the focus areas in the 5-year innovation horizon and 10-year perspective are introduced.

The final agenda builds on but does not include the full background of the extensive preliminary version (ARIADNE 2015b). For more background and evidence for the suggested innovation actions the preliminary report is available on the project website. The present report, however, includes the current ARIADNE sustainability plan, which is briefly summarised in the final section of the introduction.

1.1 Objectives, stakeholders and ARIADNE's role

Objectives

The objectives of the Innovation Agenda and Action Plan are directly related to the overall goals of the ARIADNE project: The project addresses the fragmentation of archaeological datasets in Europe by providing e-infrastructure and services that allow discovery and re-use of data held by different accessible and interoperable digital archives and collections. Thereby ARIADNE supports a culture of sharing and collaborative use of archaeological data across national, institutional and disciplinary boundaries.

In this context, the main objectives of the Innovation Agenda and Action Plan are:

- To point out innovation needs in archaeology with regard to open data sharing, digital archives and research e-infrastructure and services;
- To describe innovation potential within 5/10-year horizons, especially transformative change in research and data management practices enabled or triggered by ARIADNE;
- To suggest actions ARIADNE and other sector stakeholders can take to meet the perceived innovation needs and enable favourable changes in research practices.

It is understood that innovation in the complex fields addressed will require orchestrated measures of research institutes, digital archives, funding agencies and other stakeholders.

Innovation needs

The archaeological research and data management communities face several challenges. There are innovation needs with regard to:

- Implementation of policies and incentives for open sharing of data,
- Mobilisation of high-quality data from all relevant sources,
- Wider take-up of common data standards and machine-readable semantics,
- Provision of advanced services for data interoperability, discovery and access,
- Coordination among e-infrastructure and service providers in and beyond the sector,
- Development of environments and tools for innovative e-research.

A major aspect of these challenges is that archaeology is an extensive and multi-disciplinary field of research that spans several domains of the humanities and the natural sciences, cultural heritage research and public administration, academic scholarship as well as commercial services.

Stakeholders and beneficiaries

The ARIADNE Innovation Agenda and Action Plan address all stakeholders in research e-infrastructure and services for sharing and (re-)using openly accessible archaeological data. The stakeholders include:

- The ARIADNE project partners,
- Research policy makers and directors of funding agencies,
- Directors of research institutes and boards of research associations,
- Steering committees of collaborative research projects,
- Heads of research groups and individual researchers,
- University-based research libraries and institutional repositories,
- Directors of national and international digital archives,
- Cultural heritage agencies and museums,
- Domain and cross-domain e-infrastructure and service providers,
- Technology and software developers and providers.

Because of the diversity of the stakeholders, mutual understanding, trusting relations and joint capacity building are crucial for the successful development of open data resources, common e-infrastructure and services.

Most beneficiaries will be archaeological researchers and other end-users, i.e. cultural heritage administrators and interested citizens, for whom ARIADNE provides a portal for searching and accessing registered data. Also data archives will benefit from a common e-infrastructure and services, for example, through wider data access and recognition of curatorial work.

ARIADNE's roles

With the Innovation Agenda and Action Plan ARIADNE promotes coordination among many on-going but currently fragmented efforts to build e-infrastructures, data and knowledge resources for archaeological research in Europe and beyond. It is hoped that this will foster cohesion and synergies among initiatives, yield effective spending of available funds, and result in sustainable digital archives, e-infrastructures and services as needed by the research community.

How can ARIADNE help tackle the existing challenges? ARIADNE can play important roles in helping others to make a difference:

- *Energizer* – take the challenges as opportunities to innovate, suggest new approaches to archaeological data, invite organisations large and small to participate and contribute;
- *Enabler* – establish a platform for community building, foster close cooperation of researchers, data managers and technology experts, offer guides to good practice and support in capacity development;
- *Opener* – promote open access data, help open up 'data silos', and make shared data part of the scholarly record – persistent, available, citable and rewarded;

- *Integrator* – align currently dispersed initiatives, provide interoperability services, help to embed the use of e-infrastructure into research culture.

Stakeholders may ask if the ARIADNE e-infrastructure and services are intended to replace or complement existing facilities and services. ARIADNE does not replace existing research e-infrastructures (e.g. digital archives and preservation services) but provides integrating functionality and services on top of them. Thereby ARIADNE helps to make currently isolated archaeological data more accessible and useable. However, the ARIADNE e-infrastructure and services can only come into play when appropriate local information technology, data resources and principles of open data sharing are in place.

Stakeholders can take the ARIADNE initiative as an opportunity to organise and align their own efforts in open data mobilisation and provision. This may be stock taking (e.g. what data resources are available, how can they be made accessible, etc.), small-scale pilot projects, and participation in national and international initiatives. In particular we see good opportunities for community-building around national and other community data archives. Where digital archives are currently missing in European countries, acknowledged data centres in other countries can serve as role models.

Because of mandates by research funders to deposit and share the data produced in publicly funded research, there is an increasing demand for state-of-the-art archives. Hence there is also demand for skilled data managers at all levels, including local/institutional repositories and community archives, as well as e-infrastructures and service providers.

1.2 The preliminary report as background

The Preliminary Innovation Agenda and Action Plan has been issued about a year ago (ARIADNE 2015b). The present agenda and action plan comprises of the updated summary report of the preliminary version and includes a chapter on the current ARIADNE sustainability plan.

The extensive preliminary report (385 pages) provides the background and evidence for the actions suggested in the different focus areas of the 5 and 10 year innovation horizons.

Revisions of the summary report concern mainly the focus area 5 (Providing services and enabling novel applications) and the presentation of the 10-year horizon. The focus area 5 addresses the suggested portfolio of the ARIADNE data infrastructure services and their role in the wider information ecosystem of archaeology and beyond. The final version takes account of the current status of the data infrastructure services. The 10-year horizon section has been shortened to present its background more concisely.

As an important development since the preliminary version we note that in April 2016 the European Commission launched the European Open Science Cloud (EOSC) initiative; the EOSC will allow a tighter integration of e-research infrastructures and services (European Commission 2016).

Furthermore, the ESFRI Roadmap 2016 recognises ARIADNE as the key integrator of archaeological research data in Europe, and includes the European Research Infrastructure for Heritage Science (E-RIHS) as a new initiative on the Roadmap (ESFRI 2016: 52 and 175).

1.3 Promoting innovation over the next 5-10 years

The ARIADNE Innovation Agenda and Action Plan points out innovation needs of the archaeological research and data management community in Europe in the next 5-10 years, and suggests actions for ARIADNE and other stakeholders on how to meet these needs.

The ARIADNE initiative provides e-infrastructure and services for data networking, discovery and access. Therefore the agenda and action plan for the 5-year horizon focuses on open data sharing, data interoperability, access and re-use in further research. The 10-year horizon looks toward potential innovations in ICT-enabled archaeological research (“e-archaeology”), partly based on the progress in open data sharing and access in the next few years. The approach is summarised in the this schematic overview:

<i>Horizon</i>	<i>Innovation focus</i>	<i>Examples of suggested actions</i>
5 years (2021)	Open data sharing, data interoperability, and data access e-infrastructure & services	<i>Foster consistent data citation so that data sharers can be recognized and rewarded</i>
10 years (2026)	Novel ICT-enabled archaeological research (“e-archaeology”), partly based on the progress achieved within the next 5 years	<i>Explore virtual research environments for archaeological research</i>

The ARIADNE Innovation Agenda and Action Plan do not promote a “technology-push” view of innovation and take-up of new digital practices in archaeological research. But Information and Communication Technology (ICT), systems/tools and services play an important role in enabling novel archaeological research methods and practices. Examples of impact and success achieved in recent decades include for instance the wide use of Geographic Information Systems/Services (GIS) and 3D methods in the representation and analysis of archaeological sites.

Instead of a technology-push perspective, ARIADNE focuses on the middle ground between progress in data standards and technologies and potential users’ readiness to adopt and benefit from these advances (“adopters pull”). This may also require exploring if solutions that have been proven to work in principle (e.g. small-scale data integration projects) can be implemented and sustained at a larger scale.

5-year innovation horizon

The 5-year innovation horizon concerns immediate needs, recommended actions and measures that should be taken in order to achieve progress in priority areas, with regard to open data and data interoperability, for instance.

ARIADNE develops data infrastructure and services for the archaeological research and data management communities, e.g. dataset registration, search, access and other services. The project is not charged to provide distributed computing solutions or virtual research environments for specific domains of archaeological research. There are innovation needs in these areas, but the development and take-up of solutions will more likely fall in the 10-year horizon.

The immediate innovation needs are in the area of data sharing and access, and more of a social/institutional than technical character. Many archaeologists, like the researchers in most other disciplines, are not yet prepared to make data openly available to others outside a research project or organisation. Therefore it is vital that innovation actions contribute to the emergence of a culture of open sharing of archaeological data, trusted data archives (where missing at present), and

mobilisation of data resources that are interoperable and re-usable. The focus areas of the 5-year innovation agenda are:

- *Focus area 1 – Research e-infrastructures and digital resources:* data infrastructure for archaeological research, including coordination with infrastructures of related domains; mobilisation of high-quality data from relevant sources;
- *Focus area 2 – Culture of open sharing and re-use of data:* open data policies/mandates, removal of barriers, and promotion of data sharing, re-use and citation;
- *Focus area 3 – Data archives and curation of archaeological research data:* reliable and cost-effective community archives for long-term data curation and access;
- *Focus area 4 – Capacity building for open data sharing:* institutional policies, guidance, training and other support for open data practices;
- *Focus area 5 – Providing services and enabling novel applications:* provision of data services (e.g. dataset registration, cross-archive search and other services), and support for further novel applications for archaeological research.

The primary responsibility and agency with regard to the focus areas 2–4 is with research funding agencies, research institutions and associations, and mandated institutional and community data archives.

ARIADNE can contribute to these goals, for example, by promoting open access to archaeological data, offering guides to good practice and supporting capacity building. Many of the actions suggested for innovation through community-level data infrastructure and services in the focus areas 1 and 5 fall in the remit of ARIADNE.

10-year innovation horizon

The 10-year horizon concerns innovation potentials and developments of digital archaeology until 2026. Assuming that the 5-year objectives are met sufficiently until 2021, further progress in digital, ICT-enabled archaeological research can be foreseen, explored and targeted.

The focus shifts from the immediate need of fostering open access to re-usable and interoperable data, to enabling novel forms of archaeological e-research (“e-archaeology”). Emerging new perspectives and capabilities can be explored and pathways towards innovative and potentially transformative “e-archaeology” suggested.

In comparison to the rather clear objectives and requirements of the 5-year horizon, the 10-year horizon is a more open field. The drive towards innovative e-research in archaeology will come less from research policies (as in the case of open data), rather visionary research associations, centres and groups, data scientists, tool developers and others will pave the way towards further achievements.

1.4 ARIADNE sustainability planning

The formal lifecycle of the European Union funded activities of the ARIADNE initiative will end in January 2017. Therefore this report includes a chapter on the current status of the ARIADNE sustainability planning.

The three main lines of ARIADNE activities have been Networking/Community building, Research & technological development, and Transnational access to data services and expertise/training. The chapter discusses the possibility to maintain these activities, especially the services of the data infrastructure. A short-term sustainability solution for the core data services (dataset registry and

access portal) and pool of datasets as well as medium to long-term extension scenarios are presented and assessed.

Finally, the current sustainability plan is presented. The plan includes setting up a not-for-profit ARIADNE association for further community networking and training, and continued operation of the current dataset registry and access portal. It is assumed that sustainability of the data services will require extension, to keep the momentum (i.e. incorporation of new datasets of interested institutions), and to realise the full potential of the implemented system for the archaeological research community.

The 5-year horizon of the innovation agenda and action plan includes suggestions for stakeholders on how to support the sustainability and extension of the ARIADNE initiative (see [Section 3.2.1](#)). With regard to the 10-year horizon, the ARIADNE sustainability plan relates to the expected expansion of open access data resources and their semantic interoperability for advanced research services.

2 Overview of the innovation agenda and suggested actions

The Innovation Agenda and Action Plan points out innovation needs and challenges of the archaeological research and data management community in Europe in the next 5-10 years, and suggests actions for stakeholders on how to address the challenges. The aim of the suggested actions is to achieve favourable and potentially transformative advances in data sharing, access and re-use across institutional, national and disciplinary boundaries. This chapter gives an overview of the agenda and suggested actions for the 5/10-year innovation horizons.

In the 5-year horizon, the agenda emphasises two general, cross-cutting requirements for enabling e-infrastructures to promote innovation in research. These requirements are capacity building in open sharing of research data, and coordination of stakeholders across the different levels of data infrastructures and services. The primary objective of ARIADNE has been to develop an e-infrastructure which allows archaeological institutions and projects to share, access and re-use data resources at the community-level. Therefore, also mobilization of data/content from all relevant sources and building sustainable digital archives are major topics of the agenda.

The 10-year horizon of the innovation agenda builds on and extends the short to medium-term actions. The overall focus of this horizon is innovative digital, ICT-enabled archaeological research and communication. For this part of the agenda several future-oriented topics have been investigated, and actions are suggested which stakeholders could take already in the next few years to promote progress in digital archaeology.

The 5/10-year horizons go beyond the formal lifecycle of the current ARIADNE project. But the innovation agenda and suggested actions may help stakeholders to tackle different challenges of digital archaeology for years to come.

2.1 General, cross-cutting requirements

The innovation agenda highlights two general challenges of initiatives for community-level research data infrastructure and services.

Capacity building in open sharing of research data

Data infrastructure and services are being built to enable discovery, access and re-use of data. The investment makes little sense without open sharing of data through digital archives by the research community. Many studies on document and data repositories have shown that a “build it and they will come” approach does not work, the usage is often frustratingly small. One reason is that dedicated training and support of users is lagging behind the implementation of the technical infrastructure. There is also a need for skilled data managers at all levels, research projects, institutional and community repositories, and research e-infrastructure and services. The challenge of building and retaining a workforce of research data curators should not be underestimated.

Coordination across the different levels of e-infrastructure

Coordination of efforts is required across all levels of data management, from local efforts by university departments/institutes and research libraries, community-level national data archives, to major, European/international e-infrastructures. The primary challenges in building a rich, coherent and sustainable ecosystem of e-infrastructure and services do not concern technology, but coordination amongst the actors, building capacity and trust, managing legal aspects and, of course,

costs. The involvement of very different stakeholders and resources (e.g. research institutes and laboratories, cultural heritage agencies, museums and archives), and generally tight funds, necessitate cost-effective solutions that allow benefits at the community-level. Proliferation of uncoordinated and very likely unsustainable efforts should be prevented.

2.2 5-year innovation agenda and suggested actions

The 5-year innovation agenda comprises of five focus areas. Many of the actions suggested for innovation through community-level data infrastructure and services (focus areas 1 and 5) fall in the remit of ARIADNE. All data holders and service providers are addressed in actions suggested for the mobilisation of content/data. The research funding agencies, research institutions and associations, and institutional and community-level repositories bear the primary responsibility and agency for growing the culture and capability of open data sharing through accessible archives/repositories (focus areas 2–4).

Focus area 1 – Research e-infrastructures and digital resources

This focus area addresses three fields of activities which aim to foster sustainable and coordinated e-infrastructures for archaeological and other heritage research disciplines, take account of their domain-specific requirements, and promote the mobilisation of high-quality data from all relevant sources.

E-infrastructure development and cooperation

E-infrastructures and services support scientific activities which have become increasingly collaborative, distributed and data-intensive. In the current drive to build new or upgrade existing infrastructures and resources for e-research/science there is a risk of limited funds being invested in many uncoordinated, possibly redundant and unsustainable initiatives. This risk is particularly critical in the field of archaeological and other cultural heritage, which is characterised by having many different stakeholders and data resources of research institutes and laboratories, cultural heritage agencies, and institutions such as museums and archives. ARIADNE has developed and aims to sustain core e-infrastructure and services for archaeological research in Europe, and to strengthen the cooperation with other e-infrastructures both within and beyond the humanities and heritage sciences.

Suggested actions

- *Recognise the importance of e-infrastructure development and coordination.*
- *Support the ARIADNE data infrastructure and portal and help mobilize data providers.*
- *Ensure sustainability of the ARIADNE data infrastructure and portal.*
- *Strengthen cooperation among e-infrastructures for the humanities and heritage sciences.*
- *Establish cooperation with other e-infrastructures and services that cover relevant data of related research domains.*
- *Link up with providers of Distributed Computing Infrastructure and encourage use of their resources by archaeologists.*

E-infrastructures and novel digital practices

Research e-infrastructures can play a significant role in enhancing established forms of research as well as bringing about transformative innovation in practices and methods, in particular, in digital,

web-based environments. Several general as well as domain-specific factors and requirements must be addressed so that e-infrastructures can promote innovation in archaeology and the humanities in general.

Suggested actions

- *Take account of disciplinary differences and foster the common ground.*
- *Address the complex case of archaeology, which is a multi-disciplinary field of research.*
- *Embed e-infrastructure and services in research practice.*
- *Help overcome the current barriers to adoption and support.*
- *Build and retain a skilled workforce of data curators.*

Digital resources and services

The current landscape of content/data resources and services for archaeological and cultural heritage research is highly fragmented. Many resources are not easy to find and difficult to access or not accessible at all. A more targeted approach of resource development and access is required. The suggested actions for content/data holders and service providers are:

Register data archives, digital repositories and collections

- *Register data archives/repositories and other collections for archaeological research in the catalogue of the ARIADNE data portal.*
- *The ARIADNE registry should become the prime location to document and discover available archaeological data resources.*

Digitise cultural heritage content and make it discoverable

- *Step up the digitisation and documentation of archaeological material held by museums and other collections (e.g. unpublished excavations, grey literature, finds/objects).*
- *Implement mechanisms that allow harvesting and other programmatic access to collection metadata (e.g. OAI-based harvesting, SPARQL for Linked Data).*
- *Participate in the ARIADNE initiative for federated search and access specifically for archaeological content/data collections.*

Develop online scientific reference collections

- *Take stock of existing reference collections for archaeological purposes and evaluate their relevance and current condition (e.g. actual demand, requirements for online access).*
- *Secure financial support and curatorial expertise for the development and maintenance of state-of-the-art web-based reference resources.*

Mobilise documentation of archaeometry laboratories

- *Investigate how major facilities as well as typical archaeometry laboratories could be mobilised to support open access data (e.g. factors impeding open access and how they can be removed).*
- *Improve the online documentation of archaeometry analyses on laboratory websites, and include pointers to publications and deposited data.*
- *Standardise the metadata of archaeometry documentation to promote consistent cataloguing and improve cross-laboratory discovery of and access to publications and data.*

Virtual research environments (VREs) and data processing services

- *Investigate if archaeologists need virtual research environments; e.g. what kind of research could archaeologists conduct online, what functionalities are necessary for such research, and how could they be provided in a VRE?*
- *Promote the development of relevant VREs with functionalities (tools, services) required by archaeologists to conduct research tasks online (e-research).*
- *Raise awareness of Distributed Computing Infrastructures (e.g. Grid/Cloud-based services of DCIs) amongst archaeologists who need data processing services.*
- *Encourage use of DCIs by bringing together archaeological research groups with providers of DCI, especially with their developers of data processing software and services.*

Focus area 2 – Culture of open sharing and re-use of data

There are many good arguments for open data such as preventing duplicative data collection, allowing replication of data analysis to scrutinize knowledge claims, and enabling new research questions to be addressed based on shared data. Data that is being shared, curated, integrated, re-used and analysed with new methods gains in value rather than just being a cost factor. But there are many factors which work against open sharing of data, in particular the currently limited academic recognition and reward for making data available. All stakeholders should support open data policies/mandates and help ensure that data sharers are recognised and rewarded, e.g. data sharing could become a criterion for academic promotion and awarding of research grants. Specific measures are also required to foster openly licensed data resources, data re-use and consistent data citation. In general, archaeological data archives/repositories and data infrastructure will flourish only within a research culture that values open sharing of data.

Suggested actions

- *Promote a research culture of open sharing of data.*
- *Support open data policies of funding bodies and institutions.*
- *Suggest appropriate guidelines for open archaeological data.*
- *Recommend use of open licenses for data, metadata and knowledge organization systems.*
- *Foster consistent data citation so that data sharers can be recognized and rewarded.*
- *Conduct studies of data re-use to better understand and support current and emerging practices.*
- *Promote data re-use and highlight inspiring examples.*

Focus area 3 – Data archives and curation of archaeological research data

Data archives provide the core services of long-term preservation/curation and access to research data. They participate in e-infrastructures which enable cross-archive data discovery and access, i.e. the ARIADNE e-infrastructure for archaeological data. State-of-the-art, certified and mandated community archives allow researchers to publish their data in a secure and trusted way. Domain-based archives are the most effective solution to overcome, or at least reduce, fragmentation and inaccessibility of archaeological data resources. The costs of post-project curation and online publication of archaeological data are only a small fraction of the total project costs, around 1-3%, depending on the type of investigation and data generated. Compared to the benefits expected from

open and re-useable research data this investment seems well spent. But community archives should be stable in the long term and hence sustained support by the main funding bodies is required.

Suggested actions

- *Recognise that the costs of opening up archaeological research datasets are marginal and well spent.*
- *Include the costs of open data sharing and digital archiving in project grants.*
- *Recognise the advantages of domain-based community archives.*
- *Ensure long-term sustainability of trustworthy data archives.*
- *Encourage and support initiatives for data archives in countries where these are currently lacking for archaeologists.*

Focus area 4 – Capacity building for open data sharing

To promote the open data agenda, universities, research institutes and other stakeholders should put in place policies, guidance and training. Institutional capacity building and support for researchers in the management of data is necessary so that open and re-useable data flows into data archives for long-term curation, access and re-use. The research community should also consider novel approaches to data description and review.

Suggested actions

- *Ensure that adequate institutional policies, guidance and other support are in place.*
- *Step up capacity building and training for data management and sharing.*
- *Provide support for managing data during project work.*
- *Recognise high-quality metadata is required for data re-use..*
- *Promote data papers for archaeological datasets.*
- *Explore novel approaches to data peer review.*

Focus area 5 – Providing services and enabling novel applications

The actions suggested for the ARIADNE e-infrastructure services in the Preliminary innovation agenda concerned two main goals: to provide the services for the archaeological research community identified in the user needs surveys, and to enable the e-infrastructure act as a node in the wider information ecosystem of archaeology and beyond. The Preliminary innovation agenda was elaborated in the years 2014/2015 and meanwhile the first goal has been achieved. Therefore only the suggested actions for the second goal are retained. However, a summary of the recommended and achieved ARIADNE service portfolio is included below.

Core and additional data services

The ARIADNE user needs surveys identified the core services which the archaeological research community expects from the project. These services are:

- *A data portal that allows an overview of available but dispersed archaeological data resources;*
- *Capability to search across different digital archives/repositories which hold such resources (i.e. data collections, databases, datasets of projects, etc.);*

- *Effective data discovery, browsing and filtering mechanisms, in particular based on geo-location (maps) and date-ranges/chronologies, but also other advanced options such as faceted search;*
- *Data access methods according to the different access levels, data types/products and interaction modes offered by the providers.*

These requirements have been met by providing the user community with an online facility to register and describe accessible data resources based on a common model, and offering a portal with the desired set of data discovery and access functionalities. In addition, the ARIADNE portal provides access to special services for which a strong demand has been expressed by archaeological researchers (e.g. in project workshops). This concerns the offered services for easy web-based publication, visualization and exploration of visual media such as 3D models of objects, built structures and landscapes.

Data integration and services beyond the ARIADNE portal

The ARIADNE data infrastructure and services help overcome current limitations of discovery, access and re-use of available archaeological data. They can also enable generation of added value in the wider information ecosystem of archaeology and beyond. This can be achieved by adopting the Linked Open Data (LOD) approach to interlink ARIADNE-mediated data with various other data resources in the so called LOD Cloud. The suggested actions below comprise of two recommendations which concern such interlinking, and two others which do not necessarily require a Linked Data approach. But they may also allow generating significant added value within and beyond the ARIADNE portal.

Suggested actions

- *Deploy Linked Open Data (LOD) to integrate information within the portal and to link to external resources which follow LOD principles.*
- *Provide interfaces that allow external applications exploit available data, metadata and conceptual knowledge.*
- *Enable integrated access to data and publications, e.g. include metadata of document archives and publishers.*
- *Help enrich community websites with information about relevant data, e.g. RSS feeds on updated or newly available datasets for research on particular subjects, regions or periods.*

2.3 10-year innovation agenda and suggested actions

The 10-year innovation agenda addresses a number of topics which go beyond the immediate innovation needs such as capacity building for open data sharing and re-use based on digital archives and e-infrastructure services. Assuming that the current needs are met sufficiently in the medium-term (5 years), further progress in digital, ICT-enabled archaeological research and communication can be foreseen, explored and targeted. Therefore several future-oriented topics have been investigated, and actions are suggested, which stakeholders could take in the next few years to bring about advances in digital archaeology.

The first point on the 10-year agenda does not address digital archaeology specifically. It reflects the increasing concern of archaeologists about the societal relevance of the discipline. This issue will hardly go away, rather become more pressing. It is not primarily about better “marketing” or involving citizens, but relevant scientific contributions to societal issues. We assume that digital archaeology will be helpful to generate such contributions.

Address the question of archaeology's societal relevance

- *Leverage archaeology's societal role and relevance, for example through connecting archaeological research and knowledge with critical issues such as climate change, environmental sustainability, urban agglomeration, globalization and geo-politics, regional conflicts, migration, and others.*
- *Consider contributions which could allow archaeology a stronger voice in current debates.*

Foster participatory and reflective public/community archaeology online

- *Conceive and engage in participatory approaches based on online platforms (e.g. social media) in a highly reflective way.*
- *Explore the concept and practicalities of "open research communities" that involve archaeologists and citizens in the production, dissemination and re-use of open data.*
- *Instead of seeking "roots", focus on relevance of archaeology for current societal, environmental and other issues, at regional as well as global scale.*

Take account of the diversity of archaeological research practices and methods

- *Recognise that different archaeological schools of thought and research practices require different digital, ICT-based research environments and tools.*
- *Focus on phases in the lifecycle of archaeological research in which significant progress in knowledge may be achieved. In the last decades data generation has seen enormous progress, in the future other phases may require more attention.*
- *Recognise that issues of standardisation (e.g. data models) and cost-effectiveness are also relevant for future research practices.*

Target data integration for comparative and synthetic research

- *Foster the development of novel methods and tools that allow researchers bring together and work with the variety of data required for cross-domain, interdisciplinary research.*
- *Promote competence centres and programmes aimed to integrate data for comparative and synthetic archaeological research.*

Evolve a Web of archaeological Linked Open Data (LOD) for research

- *Promote publication of LOD datasets (collections, databases) by more archaeological and other cultural heritage institutions, especially based on mappings to the extended CIDOC-CRM.*
- *Foster a community of LOD curators who ensure reliable availability and interlinking of LOD resources (datasets and vocabularies).*
- *Develop LOD-based applications that demonstrate advances in research capability, which may motivate a wider adoption of the LOD approach by research institutions and projects.*

Explore virtual research environments (VREs) for archaeological research

- *Look into VREs developed for other domains to conceive environments relevant for e-research in specific archaeological domains as well as in cross-domain collaboration.*
- *Consider cases where researchers use data mediated by ARIADNE as well as data infrastructures and services of other disciplines (e.g. geo, environmental, biological data).*

Promote new forms of media-enriched and dynamic scientific publications

- *Promote novel forms of digital publication that “work” for archaeological projects in terms of enhanced access to research outcomes as well as academic credit.*
- *Start with moderately enriched familiar ways of publication (e.g. embedding explorable digital objects in online papers), and make new approaches as easy as possible.*
- *Investigate fields of “data-driven” archaeological research and publication in which accessible datasets and executables (software, dynamic figures, etc.) play an essential role.*
- *Explore repositories as platforms for media/data-rich archaeological publications and value-adding services for scholarly communication.*

Identify e-science practices based on data infrastructure and computing facilities

- *Promote collaborative way-finding for e-science approaches, methods and tools relevant to archaeological researchers.*
- *Focus on e-science needs specific to archaeological research, which may differ from those of other humanities as well as the natural sciences.*
- *Look for uses of low-level Grid/Cloud based services and emerging examples of archaeological applications of “big data” mining and other methods.*

Propose grand challenges for the digital archaeology community

- *Seek grand challenges that inspire the research community to push the boundaries of digital archaeology.*
- *Suggest challenges that promote mobilisation and integration of datasets for domain and cross-domain, interdisciplinary research.*
- *Bring together domain experts and developers to create methods and tools for such research.*

3 Detailed innovation agenda and action plan

The objectives of the Innovation Agenda and Action Plan are to point out innovation needs and challenges for the archaeological research and data management community in Europe in the next 5-10 years, and to suggest actions for ARIADNE and other stakeholders on how to address them. The aim of the suggested actions is to achieve favourable and potentially transformative advances in data sharing, access and re-use for further research across institutional, national and disciplinary boundaries.

The agenda and action plan for the 5-year horizon (2021) centres on challenges that are related to ARIADNE's focus on e-infrastructure and services for data networking, discovery and access. These challenges include sharing of open and re-usable data, required community data archives, and interoperability of data based on metadata and vocabulary standards. The 10-year horizon concerns potential innovations in ICT-enabled archaeological research ("e-archaeology"), which largely depend on the progress in open and interoperable data in the next few years.

We note that the 5/10-year horizons go beyond the formal duration of the current ARIADNE project (January 2017). The ARIADNE consortium has implemented an e-infrastructure for archaeological data registration, integration, discovery and access. The data infrastructure and services can support the archaeological community with regard to some of the challenges in the next 5-10 years. However most challenges need to be addressed by the wider stakeholder community, including funding bodies, universities, research institutes, professional associations, digital archives and scientific publishers. In particular these challenges concern growing a culture of open data sharing and openness of research in general.

3.1 General requirements in the 5-10 year horizons

Research e-infrastructures, digital archives and data services have become important pillars of the scientific enterprise that in recent decades has become ever more collaborative, distributed and data-intensive. There is an increasing need to enable resource sharing over e-infrastructures by pooling data resources, tools and services, aimed at supporting online team-based and cross-disciplinary collaboration. Moreover, there is the expectation that open access to re-usable and interoperable data will have an impact on the innovativeness of research, based on the exploitation of integrated datasets, for instance.

The primary objective of ARIADNE is to help overcome the existing fragmentation and inaccessibility of many archaeological data resources by making it easier to share, connect and provide access to resources at the community-level. However, the fact that today many data resources are not accessible or even discoverable online is due to data-related attitudes and practices that are not favourable for open data sharing. These attitudes and practices should change so that data becomes open, can "flow" and be re-used for further research and progress in knowledge and innovation.

Therefore, as an introduction to the innovation agenda and suggested actions, we first address two general, cross-cutting challenges for research data sharing through community archives and e-infrastructure services. These challenges entail capacity building for open sharing of research data and coordination of stakeholders at and across the different levels of e-infrastructure and services. In addition, there are discipline-specific requirements which must be fulfilled so that e-infrastructures can play a significant role in enhancing established research practices as well as bringing about transformative innovation in the humanities, and archaeology specifically.

3.1.1 Capacity building in open sharing of research data

Data infrastructure and services are being built to enable discovery, access and re-use of data. Investment in e-infrastructure makes little sense without open sharing of data through digital archives that store and curate the data, and make it available to overarching e-infrastructures and services for the research community.

The open data movement has proceeded quickly in some fields within the natural sciences, based on research culture, funder mandates, standardisation of data, and commonly available databases and archives. The situation in the humanities is more difficult and will require much capacity building. Leadership and support in the complex matters of open data sharing through accessible digital archives and other e-infrastructures is necessary.

Many studies on document and data repositories have shown that a “build it and they will come” approach does not work, the usage is often frustratingly small. One reason is that dedicated training and support of users for sharing of research results is lagging behind the implementation of the technical infrastructure. Also, changes in curricula will be necessary for the transition to open science and data practices. The responsibility for open data readiness is clearly with university departments, research institutes, and academic and professional associations.

The RECODE policy recommendations for different stakeholders on open access to research data for example urge institutions, “develop an explicit institutional research data strategy with open access as the default position”, “include open access to high quality research data as a formal criterion for career progression”, and “develop educational and training programmes for researchers and staff to improve data management skills and to enhance data-intensive research” (RECODE 2015: 8).

Indeed, open data readiness of researchers must be built at the local level, for example by departments of history and archaeology of universities, research institutes and laboratories. Moreover, there is a clear need for skilled data managers at all levels, research projects, institutional and community repositories, and research e-infrastructure and services. The challenge of building and retaining a workforce of research data curators should not be underestimated.

ARIADNE has not been specifically tasked to provide training for data curation staff. But in order to support capacity building, the project offered trans-national access for young researchers to partner facilities and expertise, i.e. summer schools and individual training (e.g. on how to make legacy data accessible). Many initiatives now offer guidance on research data management online. For archaeological data the Archaeology Data Service & Digital Antiquity guides to good practice can be recommended.

3.1.2 Coordination across the different levels of e-infrastructure

There is a need to coordinate efforts across all levels of data management, which include:

- *Local – university departments/institutes and research libraries*: provide support for individual researchers and projects (e.g. data management and access plans); strong, high-level promotion of open data is necessary at this level;
- *National – community-level data archives*: offer long-term storage, curation and access, and act as centres of expertise in data management, typically for certain domains (e.g. humanities or natural sciences); such data archives/centres are missing in many European countries;
- *European/international – major e-infrastructures*: enable transnational online access to data archives, services and tools, promote standards setting, leverage national efforts, and support

international collaborative projects; such e-infrastructures should become more integrated to enable synergies and innovative, cross-disciplinary research.

The primary challenges in building a rich, coherent and sustainable ecosystem of e-infrastructures and services do not concern technology, but coordination amongst the actors, building capacity and trust, managing legal aspects and, of course, costs.

Archaeology and other heritage sciences address a multitude of complex research topics and objects. As noted in the Roadmap 2006 of the European Strategy Forum on Research Infrastructures (ESFRI), *“The complexity of the record of human cultures – a record that is multilingual, historically specific, geographically dispersed, and often highly ambiguous in meaning – makes digitisation difficult and expensive. (...) Data, information and knowledge are scattered in space and divided by language, cultural, economic, legal, and institutional barriers”* (ESFRI 2006: 20).

Building and maintaining e-infrastructure and services for research on the record of human cultures requires coordination among all stakeholders. Lack of co-ordination would mean that available digital resources remain difficult to access and integrate, and cost-efficient operation and sustainability of the e-infrastructure and services is unlikely.

The involvement of very different stakeholders and resources (e.g. research institutes and laboratories, cultural heritage agencies, museums and archives), and generally tight funds, necessitate cost-effective solutions which allow benefits at the community-level. For example, one state-of-the-art national community archive for archaeological data is preferable to investing effort in many local solutions which are difficult to integrate and therefore not helpful in overcoming data fragmentation.

Proliferation of un-coordinated and very likely unsustainable efforts should be prevented. Integrating e-infrastructure initiatives at the national and European levels can exert a coordinative role, help prevent further fragmentation, and provide means for integration of what is recognised and actively used by the research community.

3.1.3 Overview of innovation actors and activities

Figure 1 (below) gives an overview of the many actors and activities involved in enabling to share, access and re-use research data through digital archives, e-infrastructures and virtual research environments. It distinguishes three fields of activities:

- Data mobilisation and sharing through accessible digital archives/repositories,
- Integrated data access through community data infrastructures and services, and
- Data re-use in further research, in particular collaborative research supported by e-research tools and environments.

	Data mobilisation and sharing	ARIADNE e-infrastructure and services	Collaborative data re-use
Data sharing, access and (re-)use	<ul style="list-style-type: none"> Open data policies & mandates Capacity building of institutions and researchers (training and other support) Data management & access plans Institutional repositories and community-level data archives (national, domain-based) Sharing of open, re-useable data (incl. standard metadata and common vocabularies) 	<ul style="list-style-type: none"> Networking and interoperability of data archives / repositories Data registration, catalogue of data sets and collections Data portal services (cross-archive data discovery, access and other services) Data mobilisation support (e.g. guides to good practice) Collaboration with other other e-infrastructure and service providers 	<ul style="list-style-type: none"> Re-use of shared data in new research projects (data from different institutions and countries) Virtual research communities / environments (VREs): collaborative e-research across organizations, countries, and disciplines Enhanced integrative research, comparative analysis and broad synthesis
Actors involved	<ul style="list-style-type: none"> Research funders Researchers and data managers of research projects Curators of domain archives and institutional repositories Research support services (e.g. university libraries) Academic/scientific publishers 	<ul style="list-style-type: none"> Data archives / repositories connected to community e-infrastructure Providers and managers of data infrastructure and services, ARIADNE and others Providers/curators of special resources and services (e.g. vocabularies like gazetteers, thesauri, ontologies) 	<ul style="list-style-type: none"> Research projects / collaborations Developers of open source e-research tools VRE builders and managers Distributed computing (Grid/Cloud) infrastructures and services

Figure 1: Overview of innovation actors and activities. ARIADNE 2016.

All innovation activities can be progressed in the 5-year horizon of the innovation agenda, while novel ICT-enabled collaborative e-research based on many shared data resources may see a strong take-off thereafter.

3.2 5-year innovation horizon

The objectives and recommended actions in the 5-year horizon (2021) centre on immediate challenges of open data, digital archives and e-infrastructures for data discovery and access. These innovation challenges are rather clear as they have been widely studied as well as discussed in the arenas of research policy and research infrastructure development.

The common understanding is that a research culture of open data sharing should be fostered through appropriate policies of research funders and institutions, availability of state-of-the-art data archives, and capacity building of research institutions, researchers and data managers to actually mobilise fit for re-use open data.

ARIADNE can contribute to these goals, for example, through promoting open access archaeological data, offering guides to good practice and supporting capacity development. But the primary responsibility and agency is with the research funding agencies, research institutions and associations, and mandated institutional and community data archives.

Concerning the need of an integrating e-infrastructure and data discovery and access services for the sector of archaeological research in Europe, many of the suggested innovation actions fall within the

remit of ARIADNE. The ARIADNE e-infrastructure and data portal do not replace existing data archives but provides integrating functionality and services on top of them.

The ARIADNE data portal offers a core set of required services for the archaeological research and data management communities (e.g. dataset registration, search and access services). Required related innovation objectives pertain to enabling semantic interoperability of (meta)data and knowledge organisation systems (e.g. thesauri) based on novel methods and technologies.

The sections below present the Innovation Agenda and Action Plan for the 5-year horizon. The agenda comprises of five areas, which are closely related, but can be addressed with specific actions. The suggested actions will allow enhanced access to a rich and growing stock of accessible and re-usable archaeological data.

3.2.1 Focus area 1 – Research e-infrastructures and digital resources

Background and situation

The ARIADNE project centers on providing an e-infrastructure and services for archaeological and cultural heritage data resources (e.g. data registration, discovery, access and other services). This focus area therefore addresses the current situation and major issues of e-infrastructure development in Europe in general, and e-infrastructure and digital resources for archaeological and cultural heritage research in particular.

In the last about 20 years, e-infrastructures have become ever more important, indeed crucial for the conduct and progress of research in all branches of the scientific enterprise. There is an increasing need to build, share and integrate digital resources for research. This includes terminology and conceptual knowledge (e.g. thesauri, ontologies) which allow linking of and enhanced access to research publications and data.

Archaeological research is based on established practices and tools for structuring, analysing and presenting research results of individual projects. But the development of common e-infrastructures and services for data sharing and re-use is lagging behind, although necessary to allow for more effective and possibly innovative research.

The European Archaeological Council (EAC) in the “Managing the sources of European history” theme of their Amersfoort Agenda emphasises *“the need to share, connect and provide access to archaeological information with the help of digital technologies. The key to this aspiration is to improve collaboration – we need to share rather than exchange. It is essential to encourage the development of European data-sharing networks and projects in the field of archaeology. The ARIADNE project is an excellent European initiative in this regard and participation in this project should be strongly encouraged”* (European Archaeological Council 2015: 21).

Archaeology, like other domains of research, indeed faces the challenge of supporting through common e-infrastructure and services the scientific enterprise that has become increasingly collaborative, distributed and data-intensive. But in the current drive to build new or upgrade existing infrastructures and resources for “e-science” in all research sectors there is a risk of investing limited funds in many uncoordinated, possibly redundant and unsustainable initiatives. This risk is particularly critical in the field of archaeology and cultural heritage, which is characterised by different stakeholders and data resources within research institutes and laboratories, cultural heritage agencies, and institutions such as museums and archives.

In the domain of archaeological research the risk of a “many-headed beast” scenario of e-infrastructures, initiatives heading in different directions, incompatible technologies, duplicating efforts, etc. (eResearch2020, 2010: 224-227), does not reside at the level of European-level data infra-

structure. ARIADNE has provided the archaeological sector with a common solution. But the solution will require sustained efforts to allow for maintenance and extension, mobilisation of datasets from all European Union Member States and integration with relevant other European and international data infrastructures.

Regular coordination activities at the European level are necessary to ensure interoperability of the e-infrastructures and digital resources as required by archaeological, cultural heritage and related other humanities research. Such coordination activities have already started with the humanities e-infrastructure cluster project PARTHENOS. In the medium to long term perspective also synergies with e-infrastructures of other domains of research may be sought, environmental and life sciences, for instance. Notably, we do not see archaeology as belonging exclusively to the humanities, rather this field of research is multi-disciplinary, involving research across many disciplines.

Issues of coordination and integration at the level of common e-infrastructures may be solved relatively easily. But a critical situation exists at the level of archaeological data resources. The main issues here are fragmentation, disconnectedness, and lack of open access. A more targeted approach of resource development and access needs to be conceived and put into place. But there are also some critical general as well as domain-specific factors and requirements which need to be addressed so that e-infrastructure services and tools can help drive innovation in the humanities, and archaeology specifically.

Suggested actions: E-infrastructure development and cooperation

Recognise the importance of e-infrastructure development and coordination

Research e-infrastructures and services are important pillars and drivers of collaborative and data-intensive research. They provide access to distributed, but shared digital resources (data, services, tools) for advanced and innovative research across institutional and disciplinary boundaries. Archaeology is multi-disciplinary and therefore should devote particular attention to developing integrated e-infrastructure, data resources and services.

Considerable investments have already been made at the European and national levels in implementing research e-infrastructures for different disciplines. But the e-Infrastructure Reflection Group, the main advisory body for European e-infrastructures, notes insufficient coordination among the existing e-infrastructures, and asks all stakeholders to increase coherence and resource sharing (e-IRG 2013). This is required to enable synergies, cost-effectiveness and sustainability in supporting innovative ICT-enhanced research within and across disciplines.

The sector of archaeological and other heritage research should prevent a proliferation of unsustainable e-infrastructure initiative and ensure that available funds are invested in a highly targeted and coordinated manner.

Support the ARIADNE data infrastructure and portal and help mobilise data providers

ARIADNE is the core EU Integrating Activity project in the field of archaeology, with the primary objective to help overcome the fragmentation and inaccessibility of many existing and newly generated archaeological data resources. ARIADNE is an initiative to enable this field open sharing and re-use of data as needed for progress and innovation in archaeological research, and increasingly demanded by funding bodies. The primary tasks of ARIADNE have been to set up a state-of-the-art e-infrastructure and services (e.g. data registration, discovery, access and other services) and to support data providers in connecting their resources. All stakeholders in accessible archaeological data should support the ARIADNE data infrastructure and portal and help mobilise data providers.

Ensure sustainability of the ARIADNE data infrastructure and portal

ARIADNE has developed a European-level data infrastructure and portal for archaeology. This common solution will require sustained efforts to allow maintenance and extension, e.g. mobilisation of data resources of all EU Member States and integration with relevant other European and international data infrastructures.

Strengthen cooperation among e-infrastructures for the humanities and heritage sciences

There is already a good level of information exchange between the existing European e-infrastructures for the humanities and heritage sciences, e.g. ARIADNE, CLARIN, CENDARI, DARIAH, IPERION-CH, and others. But regular cooperation targeted at enabling a coordinated development of and synergies among the e-infrastructures is necessary. Coordination is required to ensure interoperability, common policies, sharing of resources, and integrated access of users to open data, services and tools. Also cooperation on training and professional development of e-infrastructure and data service managers can be recommended.

Establish cooperation with other e-infrastructures and services that cover relevant data of related research domains

Archaeology is a multi-disciplinary field of research in which researchers need information and datasets of different domains of research. Indeed, archaeological researchers address questions about past cultures and environments that require data in the micro to macro range, i.e. from micro-remains (e.g. molecular biology data) to landscapes and beyond (e.g. terrestrial and airborne or satellite remote-sensing data). Therefore cooperation on data access between e-infrastructures of different domains is required as no single e-infrastructure may be capable to cover all data resources needed by the multi-disciplinary archaeological research communities. Cooperation on data access will also create the setting for knowledge transfer, cross-fertilization and interdisciplinary approaches.

Link up with providers of Distributed Computing Infrastructure and encourage use of their resources by archaeologists

The European Research Area has many scientific Distributed Computing Infrastructures (DCIs) and the providers welcome research communities that are not yet avid users. Archaeological research could benefit from Grid/Cloud-based services of DCIs in some cases, for example, handling and processing large stocks of data objects (“big data”), detailed 3D objects (e.g. landscapes) or complex models of cultural change (e.g. adaption to gradual or disruptive environmental changes).

Use of DCI resources can be encouraged by bringing together archaeological research groups with providers of DCI, especially with their developers of data processing software and services. We assume that currently few archaeological research groups are familiar with DCI resources and methods like science gateways, virtual machines or workflow-based execution of data processing services. Therefore guidance and active knowledge transfer by experienced developers is required.

Suggested actions: E-infrastructures and novel digital practices

Research e-infrastructures can play a significant role in enhancing established forms of carrying out research, as well as bringing about transformative innovation in practices and methods, in particular, in digital, web-based environments. But several general as well as domain-specific factors and requirements must be addressed so that e-infrastructures can become drivers of innovation in the humanities, and archaeology specifically.

Take account of disciplinary differences and foster the common ground

Instead of attempting to transfer what seemingly works in other domains (e.g. natural sciences), user-centred development of e-infrastructures, data resources and tools for archaeological research must be ensured. This requires close cooperation of researchers, data managers and technical developers in the development of innovative e-infrastructure, tools and services. The development requires fostering mutual understanding and shared perspectives. Such common ground cannot be taken for granted but must be built in close collaboration; involvement of the intended domain users will ensure that requirements are met.

Address the complex case of archaeology

Different areas on the multi-disciplinary map of archaeological research present distinctly different needs with regard to data, services or tools, classical studies versus environmental archaeology, for instance. Data infrastructures such as ARIADNE provide services required by all areas, for example data registration, discovery and access for re-use. Additional specific services/tools or virtual research environments (VREs) can be built on top of data infrastructures. They are typically area-specific, e.g. support in building a scholarly edition of classical texts or analysis of aggregated environmental archaeology datasets (e.g. vegetation distribution maps).

Embed e-infrastructure and services in research practice

Research e-infrastructures aim to empower researchers by offering them ICT-based services and tools. However, new technologies are only successful if they become embedded in the research culture and co-evolve with this culture. This requires taking account of the institutional, organisational, educational, legal and other non-technical aspects. These aspects largely determine the willingness and capability of the research community to adopt and use novel e-infrastructures, services and tools. A “technology-push” approach sidelining institutional, epistemic and methodological issues should be avoided, and conflicts due to vested interests in established research practices, technical systems and data resources must be taken seriously.

Help overcome the current barriers to adoption and support

There are several constraints which must be overcome to enable e-infrastructures to play a significant role in the humanities. These constraints pertain to lack of easy access to required data/content, effective and sustained tools, training offers and technical support. Furthermore, the priority of publications and the low recognition of digital products keep researchers disinclined to make data resources available or maintain research tools/software. Therefore the move towards open science practices will also necessitate a re-distribution of rewards, away from “high-impact papers” (or the scholarly monograph) to research databases and software. Stakeholders in open science should highlight the vital role of data(bases) and software in producing research results and ensure appropriate recognition and rewards for their sharing and maintenance.

Build and retain a skilled workforce of data curators

With regard to technical and other support we emphasise the need of skilled data managers/curators at all levels, research projects, institutional and community repositories, and research e-infrastructure and services. Data curators are part of the human element of e-infrastructures, supporting established and emerging new forms of research and discipline-specific data practices. They are also increasingly important for the success of individual research projects, e.g. with regard to the design and implementation of effective data management plans. Stakeholders at all levels (e.g. research institutions, digital archives, common e-infrastructure) must recognise the challenge of building and retaining a workforce of research data curators. This challenge should not be under-

estimated. Among other requirements appropriate recognition and clear career paths for data curators are necessary.

Suggested actions: Digital resources and services

To support the elaboration of the ARIADNE Innovation Agenda and Action Plan, we looked in greater detail into the current landscape of digital content/data resources and services for archaeological and other cultural heritage research (ARIADNE 2015a: section 9.1). The study confirmed the general perception of fragmentation and inaccessibility of many relevant data resources. This perception, and a high dissatisfaction about the situation, was expressed by researchers who participated in the ARIADNE online survey on user needs and expectations from the ARIADNE e-infrastructure and services (ARIADNE 2014: chapter 6). Many resources are indeed not easy to find (e.g. few are included in existing online registries) and difficult to access or not accessible at all. But a good understanding of the situation of different resources can allow for the conception and implementation of a more targeted approach of resource development and access.

The ARIADNE data registry and portal support this objective. But for some time to come they will hardly cover all relevant data resources for archaeologists in Europe. This is due to factors on the side of the potential providers, such as lack of technical preparedness of data resources and access mechanisms. Dedicated efforts by all data/content holders are necessary to prepare and register their resources so that they can be discovered and accessed, as well as interlinked with related information held by other institutions. Below we briefly address and suggest actions in areas of data resources and services covered in the background study. Furthermore, a summary of perceived data mobilisation priorities is included.

Registration of digital repositories / collections

In the existing registries of digital repositories (e.g. OpenDOAR, ROAR, re3data) only few European repositories and collections of archaeological data can be found. They include the ARIADNE project partners Archaeology Data Service (UK), DANS-EASY (incl. the E-Depot for Dutch Archaeology), Swedish National Data Service (incl. archaeological deposits), and the Arachne database of the German Archaeological Institute. Furthermore, there are some other, much smaller digital resources of interest to archaeology and classical studies. We think that many institutions have not yet considered registration as a way to make their collections discoverable.

Suggested actions

- *Register data archives/repositories and other collections for archaeological research in the catalogue of the ARIADNE data portal.*
- *The ARIADNE registry should become the prime location to document and discover available archaeological data resources.*

Digitised cultural heritage content

The most comprehensive overview of cultural heritage content digitisation and access in Europe has been produced by the ENUMERATE surveys (2014) which cover the situation at 1179 institutions. The surveys give some indication concerning digital/digitised information of relevant museum collections and archaeological sites, monuments and landscapes. For example, of the collections of 548 museums (including 180 “archaeology or history museums”) 24% were already digitised, 57% not yet digitised, and for 19% no need was perceived to digitise them. Collections of 239 institutions that hold mainly or a significant amount of material concerning archaeological sites, monuments and landscapes show roughly the same pattern, but a smaller percentage (18%) have already digitised

content. The ENUMERATE surveys also found that the institutions primarily provide digital access on their own website and show a low interest in allowing for programmatic access.

Suggested actions

- *Step up the digitisation and documentation of archaeological material held by museums and other collections (e.g. unpublished excavations, grey literature, finds/objects).*
- *Participate in the ARIADNE initiative for federated search and access specifically for archaeological content/data collections.*
- *Implement mechanisms that allow harvesting and other programmatic access to collection metadata (e.g. OAI-based harvesting, SPARQL for Linked Data).*

Online scientific reference collections

Scientific reference collections are difficult to maintain, re-vitalize and extend, due to shrinking budgets for their curation. In recent years these collections have been promoted as essential research infrastructures, but substantial investment will be necessary to turn many collections into state-of-the-art web-based resources. Some institutions have already digitised collections that are relevant for archaeological purposes but discovery and access is not well organised at the community level. Bottom-up initiatives of researchers aim to develop online community resources from individual contributions, but usually receive only few.

Suggested actions

- *Take stock of existing reference collections for archaeological purposes and evaluate their relevance and current condition (e.g. actual demand, requirements for online access).*
- *Secure financial support and curatorial expertise for the development and maintenance of state-of-the-art web-based reference resources.*

Documentation of laboratories for archaeometry

In the MERIL inventory 17 laboratory facilities relevant for archaeometrical analyses are registered. These are large facilities which have special equipment such as synchrotron radiation and accelerator mass spectrometry instruments, typically not available at archaeometry laboratories of university institutes and large museums (of which there exist many more). An examination of the websites of the large laboratories shows that documentation (metadata) of their analyses of archaeological and other cultural heritage material is hard to come by and access to data is offered nowhere. Better documentation could greatly improve the discovery of such analyses and finding related publications and data in repositories, which may not necessarily be at the laboratories.

Suggested actions

- *Investigate how major facilities as well as typical archaeometry laboratories could be mobilised to support open access data (e.g. factors impeding open access and how they can be removed).*
- *Improve the online documentation of archaeometry analyses on laboratory websites, and include pointers to publications and deposited data.*
- *Standardise the metadata of archaeometry documentation to promote consistent cataloguing and improve cross-laboratory discovery of and access to publications and data.*

Virtual research environments and data processing services

Virtual research environments (VREs) combine tools and services for domain-specific online research purposes (e-research). In the field of archaeological research few VREs are available or in development. Sometimes Web 2.0 applications are suggested, however, they mainly serve the professional networking and information exchange function, while other functions (e.g. data management and processing) require different systems. For data processing Grid/Cloud-based services of Distributed Computing Infrastructures (DCIs) allow user groups to share data resources and processing applications (virtual machines) and, thereby, form a virtual research community. ARIADNE does not foresee offering a DCI. A VRE could be implemented on top of or related to the ARIADNE e-infrastructure for using existing as well as creating new data resources.

Suggested actions

- *Investigate if archaeologists need virtual research environments; e.g. what kind of research could archaeologists conduct online, what functionalities are necessary for such research, and how could they be provided in a VRE?*
- *Promote the development of relevant VREs with functionalities (tools, services) required by archaeologists to conduct research tasks online (e-research).*
- *Raise awareness of Distributed Computing Infrastructures (e.g. Grid/Cloud-based services of DCIs) amongst archaeologists who need data processing services.*
- *Encourage use of DCIs by bringing together archaeological research groups with providers of DCI, especially with their developers of data processing software and services.*

Summary of data mobilisation priorities

ARIADNE provides a facility to register data resources so that what is accessible can be documented, made visible and discoverable. The ARIADNE online survey found that gaining an overview of and ability to search across available resources is on top of what researchers expect from the project. The ARIADNE data registry and portal allow this, for large and small datasets, if many data holders make the effort to register the data.

E-infrastructure initiatives like ARIADNE of course want to have a critical mass not only with regard to the number but also volume of registered resources. Therefore incorporating available archaeological data from large community-level and institutional resources is clearly among the priorities of the initiative. The overall tenor of the responses to the ARIADNE online survey (ARIADNE 2014: 76-84) was that *all* types of data are important; exemplary comments were *“all sources are important and must be first-hand”*; *“all is important ...if it is available”*. This of course is related to the existing difficulty of finding and accessing online the many data resources researchers would like to consult and use for their research.

A comparison of what respondents considered as important for their research with their rating of the online accessibility of these resources gives some general indication of data mobilisation priorities: excavations, prospection/field surveys, GIS, material/biological analyses and radiocarbon/dendrochronology showed the largest gaps between importance and actual online accessibility of data. Below we address data resources, which may be priority targets and require specific actions by data holders:

Excavations, prospection/field surveys, including GIS

Some information about excavations, prospection/field surveys, including Geographic Information System/Services (GIS) used, is usually available online from project websites. But more could be done

by research groups to make data available online during projects (which may extend over many years), and ensure that the full data/evidence of completed projects is deposited in a digital archive.

Legacy data and grey literature

Much excavation and other archaeological work has been conducted which remained unpublished and poorly documented. National initiatives to identify and unearth important unpublished work could add greatly to the accessible archaeological record.

A lot of unpublished fieldwork reports (“grey literature”) have been produced in the context of infrastructure development and other land use projects. While often the only record of the fieldwork undertaken, such reports are difficult to find and access in many European countries. Among the exceptions is the UK, where provisions have been put in place so that fieldwork reports are deposited with the Archaeology Data Service (at present about 40,000 reports). The extensive usage makes clear the need to make such material more accessible.

Laboratory analyses of archaeological and other heritage objects

Concerning material/biological analyses, our study confirms the researchers’ perception of a difficult situation with regard to discovery and access; suggested actions are included in the brief summary above (for background see ARIADNE 2015a: section 9.1.3). Concerning dendrochronology and radiocarbon data, also on the wishlist of researchers, we found several relevant databases but did not look deeper into the situation (e.g. regional coverage, ease of data access and re-use, etc.).

Non-priority data resources

The comparison of what the ARIADNE survey respondents considered as important for their research with their rating of the online accessibility of these resources shows two non-priority areas: One area is government site management data/information (e.g. sites and monuments records), although their online accessibility was perceived as low. Also not a mobilisation priority is large datasets or stocks of data (“big data”) which could be used for data mining or other advanced computing methods. Reasons may be that respondents did not expect much from such approaches in archaeology and may lack expertise to apply the required technology/software and methods. The online availability of data resources for applying such methods was perceived as very low.

3.2.2 Focus area 2 – Culture of open sharing and re-use of data

Background and situation

ARIADNE is the core EU Integrating Activity focused on archaeological data(sets). Through its e-infrastructure and services ARIADNE enables enhanced discovery, access to, and re-use of archaeological data held by institutions across Europe (and beyond). But ARIADNE, like other data infrastructure initiatives, depends on underlying institutional repositories and community archives richly filled with open data shared by research projects.

Various benefits are expected from the ability to re-use open data for further research and other purposes. Research institutes, data archives and e-infrastructure providers, funding bodies and other stakeholders therefore cannot ignore obstacles which hinder researchers in sharing their data in an open manner. Rather they must support researchers in data sharing and help ensure that they receive appropriate credit for doing so.

Arguments for open data sharing

There are many good arguments for open data sharing such as preventing duplicative data collection, allowing replication of data analysis to scrutinize knowledge claims, and enabling new research questions to be addressed, i.e. based on combined datasets. Particularly strong are the arguments

concerning the re-use of data as re-use allows exploitation of previous investment. Data that is curated, integrated, re-used and analysed with new methods gains in value rather than just being a cost factor. The benefits expected from data re-use are among the core interests of research funders in open data.

Data sharing through accessible archives

Researchers often share data, but mainly with project collaborators and other trusted colleagues. Most data remains within a small circle of peers and is not available for other researchers or the public at large. The goal of open sharing is to have the data deposited in digital archives to allow for access and re-use in further research and for other purposes. Instead of inaccessibility and potential loss, deposited data is curated and accessible for new, often not anticipated uses. However, strong barriers need to be addressed to move from a closed-circle to open sharing of data.

Barriers to open data sharing

Many factors work against open sharing of research data. In particular, researchers perceive a lack of incentives (e.g. little academic recognition and reward), and fear that their data could be misinterpreted or misused. Preparing data for re-use by others means extra work (e.g. data description), while there are other priorities such as new projects and publications. The core requirements for open data sharing are not technical but institutional, especially the need for appropriate academic recognition and reward. But also other requirements must be met, such as available, trusted and sustainable data archives. At the international level, social and technical requirements of open data sharing are being addressed by working groups of the Research Data Alliance.

Strict open data mandates

A broad consensus is emerging that research data, in particular, data from publicly funded projects should become openly accessible for all. Ever more research funders extend their open access mandates for publications to the data that underpins the research results (e.g. the European Commission and many national funding bodies). To ensure compliance, project funders also require grantees to implement a data management plan, with the objective to make the generated data publicly accessible through an appropriate repository. Experience shows that open access policies/mandates of research funders and institutions for publications and data need to be decisive, monitored and enforced. Also journals that demand research underlying reported results to be accessible have a significant impact on the availability of the data.

Data re-use and citation

Not a strong driver of open data sharing already is evidence of individual benefits for researchers. Such benefits can be accrued through re-use and citation of shared data by others. While a citation advantage for open access publications is well confirmed empirically, there is little evidence as yet of a similar advantage for open research data. Also solid empirical evidence of re-use of openly shared data is not available. But we may expect better insights into data re-use practices as more open research data becomes available. One important element in this regard is standardisation of data citation in the research literature and other publications (e.g. websites), including persistent identifiers of the data(sets).

The archaeological data sharing record

Archaeology is not mentioned above, because the issues and required solutions are the same in all domains of research. But some are further ahead concerning core requirements, decisive open access policies and available and mandated data archives, for instance. In European countries where the requirements are fulfilled, archaeology is also part of the open data movement.

Elsewhere archaeological research is not at the forefront of open data sharing, despite factors that make an open data imperative particularly strong in this field: excavation of sites destroys the primary archaeological evidence, the work on archaeological heritage is done in the public interest, and there is little commercial interest in archaeological data (apart from the commercial archaeology sector).

However, many archaeologists are not yet well prepared or equipped for open data sharing. As the matter is complex, strong leadership with regard to policies/mandates, supportive institutional measures (e.g. capacity building, training of researchers), and state-of-the-art digital archives are required. Overall this is about growing a culture of open data sharing, based on common values and agreed principles. As a popular slogan of OpenAIRE puts it, *“Open Access is global – but implementation is local”*.

Openness should become embedded in archaeological research practices as *“the default modus operandi”* (e-InfraNet 2013) so that the advantages of accessible and re-usable data gain priority over the interest of the individual researcher to protect his/her *“claim”*.

Suggested actions: Open data policies and practices

This focus area mainly concerns the challenge of growing an open culture of research of which open data is but one, although an important facet.

Promote a research culture of open sharing of data

Archaeological data archives/repositories and the ARIADNE data infrastructure will flourish only within a research culture that values open sharing of data. Therefore promoting this culture is of vital interest to data service providers which, however, have a supportive role. The main enablers are research groups, institutions, associations and funding bodies who opt for openness. There are many good arguments for open data such as preventing duplicative data collection, allowing replication of data analysis to scrutinize knowledge claims, and enabling new research questions to be addressed based on shared data. In particular, data that is shared, curated, integrated, re-used and analysed with new methods gains in value rather than just being a cost factor.

Support strict open data policies of funding bodies and institutions

Empirical studies make clear that only decisive open access policies for publications and data work effectively. Therefore it is recommended that funding bodies and research institutions issue strict mandates for open data sharing. Data archives and data infrastructure providers should give full support to such mandates.

Suggest appropriate guidelines for open archaeological data

A *“one size fits all”* approach will not work for archaeological data. In case open data guidelines do not take account of specificities of archaeological data (e.g. sensitive data such as find-spots, human remains, knowledge of indigenous peoples), research institutions and associations should promote community consultation and suggestion of appropriate guidelines. The move towards open data offers an opportunity to discuss and get clear about disclosure or non-disclosure of certain information.

Recommend use of open licenses for data, metadata and knowledge organization systems

Research funders, institutes and archives should make sure that data, metadata and knowledge organization systems (e.g. thesauri) are shared under an appropriate open license. Data licenses should not impede effective re-use and further dissemination of the data, but ensure attribution of

data sharers. As appropriate are considered Creative Commons (CC), Open Data Commons (ODC) and other licenses of type “BY” (Attribution), which ensure that data sharers (researchers, institutes and others) are recognised. For metadata the most open licenses are recommended, e.g. CC Zero (CC0), ODC-PDDL or other public domain dedication licenses. Licenses containing a No-Derivatives condition should generally be avoided, and there are strong arguments against Non-Commercial; applying instead ShareAlike will impact negatively on the ability of intermediation services to combine and utilize (meta)data effectively. Wide application of open licenses is necessary to remove barriers to legitimate and effective re-use, extension and integration of data(bases), metadata and knowledge organization systems (e.g. thesauri).

Help ensure that open data sharers are recognised and rewarded

Archaeology should be a leading example of open data sharing, because of factors that make an open data imperative particularly strong for archaeology: excavation of sites destroys the primary archaeological evidence, archaeological work is conducted in the public interest, and there is little commercial interest in archaeological data. But there are many factors which work against open sharing of archaeological data. A particularly strong factor is little academic recognition and reward for making data available. The academic credit system values research publications, while data sharing is considered much less, if at all. Therefore all stakeholders in open archaeological data should help ensure that data sharers are recognised and rewarded. Institutions that make data sharing a criterion for academic promotion, and funding agencies that take it into account when awarding new research grants will have the most impact in this regard.

Suggested actions: Data re-use, citation and altmetrics

Benefits associated with re-use of openly shared data arguably are the strongest drivers of the open data agenda. For example, re-use in further research (e.g. based on combined data) is expected to provide much return on investment. Re-use is also particularly important to individual researchers as it can bring recognition and rewards to data sharers based on data citations. Data archives and e-infrastructures could help mobilise more open research data more easily if convincing evidence for the expected individual benefits is available. However, empirical evidence for such benefits is still scarce. Existing archaeological data archives are accessed frequently, but little is known about actual data re-use and citation. Below we suggest how a better understanding of data re-use could be acquired, re-use promoted, and ensure that data sharers receive the credit they deserve.

Conduct studies of data re-use to better understand and support current and emerging practices

Surveys or a series of case studies could allow evaluation of current practices of data re-use in different areas of archaeology. Information about the main current and emerging forms, contexts and requirements of data re-use would be helpful to possibly better support these practices. In particular, data archives, e-infrastructure and tool developers could benefit from such information.

Promote data re-use and highlight inspiring examples

Research organisations, data archives/repositories and e-infrastructure service providers should promote data sharing and re-use wherever appropriate. Specific measures should be applied such as data re-use competitions and highlighting of inspiring examples. The introduction of a data re-use award by the Archaeology Data Service (UK) is a good example for this approach.

Foster consistent data citation so that data sharers can be recognised and rewarded

Researchers who re-use data from digital archives or databases should reference the data in a standardised way. The standardisation within and across disciplines is still in the making. We

recommend that data archives/repositories, journal publishers and other service providers in the field of archaeology promote a common form of data citation in the research literature and other publications (e.g. websites). Among the requirements for consistent and accurate reference of (re-) used data we highlight persistent identifiers (e.g. DOIs or other). Standardised citation can greatly enhance the identification of data re-use and recognition of researchers and data archives that make data available. Eventually it will enable tracking of data re-use and support evaluation of the impact of shared data.

Capture and present data usage figures

Data archives/repositories and researchers will for some time to come lack information if accessed research material has been re-used and cited. Therefore other indicators of usage (e.g. views, downloads) should be applied and presented. Such indicators are called alternative metrics (altmetrics) and also include social media activities around research output.

Usage figures can allow data archives and other service providers to demonstrate relevance to the research community and funders, plan further development, and mobilise new data contributions. We recommend usage indicators that are directly related to the material that is being curated and made accessible (e.g. views, bookmarks, downloads); social media activities around the content/data are not seen as valid indicators by most researchers. It is also recommended to communicate the value curation work adds to data collections (e.g. indicators of improved discovery and access).

3.2.3 Focus area 3 – Data archives and curation of archaeological research data

Background and situation

Digital archives for deposit, long-term curation of and access to data provide core services underlying research e-infrastructures. Data archives richly filled by the research community are one key to the success of e-infrastructure initiatives, because ultimately researchers want to discover and re-use available data for purposes such as comparison and further research. At the same time many researchers are reluctant to share their own data beyond project collaborators and other trusted colleagues. Therefore motivation and support of open data sharing is necessary.

State-of-the-art community data archives can foster trust in open data sharing as they provide a reliable environment for data publication, re-use and citation. We highlight national and international research community archives as they can substantially reduce data fragmentation. The “institutional” model, repositories of single universities or research institutes, provide a local solution for affiliated researchers but offer only limited potential to overcome the fragmentation of archaeological data.

Extending the productive lifecycle of data

Ensuring long-term curation of and access to archaeological datasets goes beyond the capability of individual researchers, research projects and, arguably, most institutes. In any case, the existing difficulty of discovering and accessing archaeological datasets demonstrates that individual “local” solutions are not appropriate for the task. Therefore the data should be deposited in archives that ensure long-term curation and access.

Unfortunately, the lifespan of archaeological data often ends when researchers have published their results, but do not make the underlying data available. Extending the data lifecycle through measures of curation and dissemination allows use of the data in further research, for example, exploration of integrated datasets with novel methods and tools. A core role of digital archives and e-infrastructures therefore is mobilising and bringing together data resources so that they can be used

in novel ways. The expectation is an increase in value of existing data through wider and new uses within and beyond the research community.

Towards an optimal solution

The fragmentation of archaeological data poses enormous difficulties to achieve aggregation and integration of data of many projects/institutes at the level of common e-infrastructure and services. As the most effective solution for tackling the issue we see domain-based, community-level archives. The advantages of such archives include that they can:

- Provide clear orientation for all stakeholders concerning data mobilisation, expected good practices, etc.;
- Act as centres of expertise required for the types of data and data-related issues common in the domain;
- Foster trust in open data sharing through providing a reliable environment for data publication, re-use and citation;
- Allow cost-effectiveness of data curation and access (e.g. economies of scale);
- Promote common standards and act as hubs for data integration and access.

References for such digital archives exist, for example, the Archaeology Data Service (UK, established 1996) and the E-Depot for Dutch Archaeology of the Data Archiving and Networked Services (Netherlands, since 2005); both are mandated archives for depositing data of archaeological research in the respective country (and ARIADNE partners). In France, since 2014 ARIADNE partner Inrap, the Institut National des Recherches Archéologiques Préventives, archives fieldwork reports (old and current) with the Centre Informatique Nationale de l'Enseignement Supérieur (CINES). CINES provides a data archive for long-term preservation. Also archaeological 3D projects supported by CNRS Archéovision archive their data with CINES.

Many European countries lack an appropriate digital archive where archaeologists could deposit their data. Not a good solution for overcoming data fragmentation in such cases, especially large countries, would be several research institutes each trying to establish their own data repository, or archaeological data ending up in many general-purpose university-based repositories. Therefore the IANUS Research Data Centre for Archaeology and Classical Studies in Germany (coordinated by ARIADNE partner the German Archaeological Institute) is a consequent initiative for a common solution. New initiatives in this direction have been inspired by ARIADNE in smaller countries, Austria, Ireland and Slovenia, for instance.

Mandates and certification

Research funders mandate domain-based archives that are acknowledged for their high standards of data curation or, if such an archive is not available, recommend use of a recognised and suitable other repository. Acknowledged state-of-the-art archives devote special attention to measures which promote trust and credibility. One measure is certification according to a standard of trustworthiness. Both the Archaeology Data Service (UK) and Data Archiving and Networked Services (Netherlands), and several other digital archives, are certified according to the criteria of the Data Seal of Approval.

In the absence of a decisive deposit mandate, i.e. deposit is only recommended and several archives are considered as relevant, the certification of a data archive will be one factor in researchers' selection of an archive. According to an available study archaeologists appreciate most an archive's transparency with regard to the description and management of data collections, and indications of archive stability (e.g. sustained funding).

Benefits and costs

The benefits of a community-level, national/international data archive stem from its role as a central and sizeable hub of data in the research field it supports. Reliable, one-stop access to needed information and data resources makes research easier, faster and cheaper. In the case of the Archaeology Data Service (ADS) the increase in research efficiency of the users has been calculated to be worth at least 5 times the costs of operation. Including other advantages, a £ 1 investment in ADS yields up to £ 8.30 return. Like the ADS also other richly filled digital archives may have many non-academic users, including heritage management, educational, and private/general interest user groups (e.g. local/regional history).

Concerning costs, one important fact to bear in mind is that the costs of post-project data curation and online publication of archaeological projects are only a fraction of the total project costs, between 1-3%, depending on the type of investigation and data generated.

Another rule of thumb is that running an archaeological or other data archive of course costs considerably more than a typical institutional document repository. One major cost driver in archaeology is the variety and complexity of the data that needs to be ingested and curated. Data acquisition and ingest are the most costly curation activities, while archival storage and preservation activities are a much smaller segment of the overall costs, and likely to decline over time.

Keeping the operational costs stable while curating larger data collections allows economies of scales (lower per-unit cost). Enhancement of labour-intensive curation activities through streamlining and tool-support (e.g. easy submission of small deposits) can allow significant cost reduction.

Suggested actions: Data archives and curation of archaeological research data

Digital archives for long-term curation of and access to data provide core services underlying research e-infrastructures. Archives richly filled by the archaeological research community are one key to the success of the ARIADNE e-infrastructure initiative, which provides a solution for cross-archive search, access and re-use of available data resources.

Researchers are looking for relevant data, but are not necessarily willing to share their own data in an open manner. State-of-the-art, certified community archives allow researchers to publish their data in a secure and trusted way. The publication comes with a price tag, however compared to the total costs of archaeological investigations these costs are relatively small.

Substantial benefits are expected from the availability of ever more open research data. The chances that these benefits materialise appear to be higher if mandated community-level data archives are established. In any case, such archives will reduce the fragmentation and inaccessibility of many archaeological data and provide a more cost-effective solution than spreading funds across many data repositories of individual institutes.

Recognise that the costs of opening up archaeological research datasets are marginal and well spent

The costs of post-project data curation and online publication of archaeological projects are only a small fraction of the total project costs at around 1-3%, depending on the type of investigation and data generated. Compared to the many benefits expected from open and re-usable research data this investment seems well spent.

Include the costs of open data sharing and digital archiving in project grants

All surveys on open data sharing show that researchers consider the related effort as a significant barrier. Specifically, this concerns the effort required to prepare shareable data and detailed data

description. Therefore research funders should allow inclusion of the costs of this work in project grants. A project data management plan, as increasingly requested by research funders, is the ideal place to present these costs as well as the expected archive charge for long-term data preservation and access.

Recognise the advantages of domain-based community archives

Building and mandating community data archives is the most effective strategy to overcome, or at least reduce, fragmentation and inaccessibility of archaeological data resources. Advantages of central, domain-based archives include clear orientation for all stakeholders, focused mobilisation of data deposits, economies of scale, among others. Substantial return on investment through research efficiency and use also by non-academic groups can be expected from domain archives that provide access to a wider range of relevant data resources. Funding one data centre is very likely a better solution than spreading funds across many hard to sustain data repository projects of individual institutes.

Ensure long-term sustainability of trustworthy data archives

A clear priority for open data initiatives is that archives are available where researchers can share their data in a secure and trusted way. Such archives should be certified and stable in the long-term in order to promote trust that the effort put into re-usable data is well spent. Commitment for sustained support (10+ years) by the main funding bodies would be helpful in this regard.

In the current economic climate it seems unlikely that additional funds for open data curation and access can be mobilised. This means that budgets may need to be shifted from already restrained research project funding to appropriate data curation and access solutions. Related concerns should be addressed by making the benefits of open research data clear.

Encourage and support initiatives for data archives in countries where these are currently lacking for archaeologists

The most effective solution for open archaeological data mobilisation, curation and access is a domain-based central archive (e.g. one place to go, community building, cost-effectiveness, and others). Yet such a data archive is missing in many European countries. Progress towards common solutions may be promoted through knowledge transfer between established data centres and initiatives for new archives in other countries. New entries may “leapfrog” to a state-of-the-art solution by learning from acknowledged benchmarks.

Devise interim archiving solutions for datasets that are inaccessible and at risk of loss

Archaeological research institutions should take stock of valuable old datasets as well as datasets not maintained by researchers actively using the data. Where a data archive for long-term curation and access is missing, devise an interim local or outsourced solution to prevent inaccessibility and potential loss of valuable datasets. A priority of course is excavation data that cannot be re-collected.

3.2.4 Focus area 4 – Capacity building for open data sharing

Background and situation

This focus area addresses the need of capacity building for open data sharing at the institutional level and issues concerning the required quality of data description and review. To support research project data management and sharing, universities and research institutes will have to put in place adequate institutional policies, guidance, training and other support. Strong institutional support of the open data agenda will allow more archaeological data (and metadata) flow into data archives for long-term curation, discovery and access, and re-use in further research.

Current situation of institutional data management support

At present the level of data management support by libraries/repositories of universities and other centres of research and education is rather low. The focus is mainly on institutional policy support and advice (e.g. issues of open access, data management plans, IPR/licensing, etc.), and organising training for research students. The main challenges for the research libraries, and institutional document repositories they manage, are limited resources (funds, personnel) and skills gaps with regard to research data curation. Moreover their potential role in managing or providing support in the management of research data is an on-going topic of discussion.

There are considerable doubts about attempts of individual universities to implement research data curation. Proper curation of research data requires specialisation, and leading examples of data archiving services indeed are specialising in research fields. Examples are the data centres funded and mandated by the UK Research Councils (e.g. Archaeology Data Service), or Data Archiving and Networked Services (Netherlands), which centre on social sciences and humanities (including archaeology). Specialisation is difficult to achieve for many different disciplines, e.g. all disciplines present at a university. Also a small or medium-size archaeology institute will hardly be in a position to maintain a data archive, and generally prefer to focus on research rather than data curation.

Support for managing data during project work

Rather than building research data archives, it is preferable that university departments and research institutes focus their efforts on training and support for researchers, so that shareable data emerge from the research process and are provided to appropriate subject/domain-based archives. More efforts could also be devoted to making legacy data, reference collections and other institutional assets available.

It is understood that proper research data management should start and be supported as early as possible, while researchers are working on their projects. Data management plans as requested by ever more research funders can provide a basis for this support. For the management of “active data” various methods are suggested, which range from plug-ins for software already used by researchers to automatic capture of information (metadata), which would reduce the effort required for regular documentation of research activities and results. At a minimum institutions or research networks could offer researchers a safe and controlled environment for storing and sharing data during collaborative project work.

High-quality metadata as required for data re-use

When researchers go about sharing data through an archive/repository the question of metadata comes up. All studies on data sharing through digital repositories (including the ARIADNE survey) found that researchers consider the effort to provide the required metadata as a barrier to open data sharing. While data repositories and users would benefit from high-quality metadata, data sharers face the burden and usually prefer not to invest much effort on providing metadata.

Asking for high-quality metadata is likely to result in fewer contributions, which is one reason why many repositories have rather shallow discovery metadata. Such metadata is insufficient to assist data re-use, which requires data description so that potential re-users can understand the data provenance/context, evaluate if the data is relevant for intended purposes, and use it properly to prevent incorrect conclusions. Requesting high-quality metadata arguably is possible only for domain-based archives that are mandated or recognised as the best place to share valuable data according to community standards.

Data papers and data review

A novel approach that offers an incentive for researchers to provide rich data description is the peer-reviewed “data paper”. Ever more journals are established that invite researchers to describe datasets they have deposited or databases they maintain. Moreover data review has become a topic of much recent discussion. It is anticipated that the open data policies of research funders and journals will bring about a wave of data in need of quality review. Therefore new models of peer-review need to be considered.

Suggested actions: Capacity building for open data sharing

To promote the open data agenda, universities, research institutes and other stakeholders should put in place policies, guidance and training. Institutional capacity building and support for researchers in the management of data is necessary so that open and re-usable data flows into data archives for long-term curation, access and re-use. The research community should also consider novel approaches to data description and review.

Ensure that adequate institutional policies, guidance and other support are in place

A growing number of funding bodies require research projects to implement a data management plan, with the objective that the generated data becomes openly available. Therefore researchers are looking for advice, guidance and support from their institution. Many institutions may not be prepared to meet this demand, because the responsibility to manage and maintain data has traditionally been assigned to the researchers, based on the understanding that they own the data. We recommend that research institutions address issues of proper data management and sharing pro-actively and put adequate policies, guidance and other support in place.

Step up capacity building and training for data management and sharing

There is no lack of guidance material for good practice data management. For archaeological projects particularly the guides offered online by Archaeology Data Service & Digital Antiquity and the ARCHES guide for data archiving (available in several European languages) merit highlighting. The challenge for publicly funded research projects and institutions now is to implement and support data management in view of open data sharing. This necessitates institutional capacity building, training offers and other measures aimed at ensuring data stewardship and accountability.

In practical terms, this means a range of assistance including, but not limited to, hands-on training in data management for PhD students and early-career researchers, help in drawing up a solid data management plan for discipline-specific data, expert advice concerning sensitive data and ethical issues, IPR/licensing (incl. rights clearance). Addressing “the data issue” therefore means more than pointing researchers to an appropriate data archive, institutional capacity building is necessary to provide the kind of assistance mentioned.

Below we give some specific suggestions on how research institutions and archives can foster proper data management and sharing:

- *Require that data management and subsequent open sharing of data are considered already in the project planning phase, e.g. data management and access plans.*
- *Promote the preparation of shareable data through dissemination of good practice guides and expert advice on specific matters (e.g. sensitive data, licensing of databases).*
- *Bring project data managers in contact with data archives/centres, not only in view of data deposit, but skills development (e.g. work on legacy data, use of specific standards).*
- *Emphasize the need of appropriate description of the methods used to collect, analyse and present the data, including technical and other requirements for data re-use (e.g. software).*
- *Suggest use of established open data formats, metadata standards as well as common terminology/vocabularies.*

Provide support for managing data during project work

Advocates of proper research data management suggest that it should start and be supported professionally as early as possible. This would make on-going research work more effective and data re-use easier. Suggested approaches include offering researchers lightweight curation tools and/or automated mechanisms integrated into researchers' normal workflows (so called "sheer curation"). Solutions ideally add immediate value to the creators and primary users of the data as well as prepare the ground for long-term data archiving and access.

This is still a field of research & development with some prototypic solutions. Also use of workflow management systems, for example, in archaeometry laboratories seems uncommon. Rather than looking for automation, we suggest that institutions or research networks offer researchers a safe and controlled environment for storing and sharing "active data" during project work. Moreover, it would be beneficial if institutional data experts are available who support on-going projects, with regard to metadata generation and use of common vocabularies, for instance.

Recognise high-quality metadata is required for data re-use

Research datasets that are shared through repositories must be provided together with metadata, which are necessary for data discovery and access. But discovery metadata may not be sufficient for data re-use. Data description is required so that potential re-users can understand the data provenance/context, evaluate if the data is relevant for intended purposes, and use it properly to prevent incorrect conclusions. The effort to produce such data description is among the top barriers for open data sharing.

Metadata presents a conflict of interest between repositories, who would like to have good metadata, and data depositors, who may not be willing to invest much effort on the task. In practice this means that there are many repositories which have only shallow metadata, in particular general-purpose and university-based repositories. Both invite deposits of material from many disciplines hence cannot support specific domains like archaeology specifically. Domain-based, specialised and mandated archives can set high-quality metadata standards, which depositors will accept and follow, guided by archive curators, if necessary.

Promote data papers for archaeological datasets

A data paper is a peer-reviewed publication that describes a data resource (dataset, database or other), the methods and standards used to create it, its structure and size, and where and how it can be accessed. In particular, the paper should also describe the re-use potential of the data for further

research or other purposes. Peer-reviewed and citable data papers offer an incentive for researchers to make their data available through repositories and provide rich description of the data intended to promote and enable re-use. One example of a dedicated, online and open access data journal is the *Journal of Open Archaeology Data*; the e-journal *Internet Archaeology* has initiated a series of data papers, and others may follow suit.

Explore novel approaches to data peer review

The open data policies of research funders and journals have raised the question of how the quality of the data should be evaluated. Peer reviewed data papers provide one solution, furthermore novel models recently developed for publications may be trialled. Such models are editor-mediated review, “crowd-sourced” review and post-publication review. For example, Data Archiving and Networked Services (DANS) collect ratings and other feedback from people who have downloaded data from their archive system, which includes the E-Depot for Dutch Archaeology.

3.2.5 Focus area 5 – Providing services and enabling novel applications

Background and situation

Archaeological research in Europe (and elsewhere) needs common and integrated e-infrastructure and services for data curation, discovery, access and re-use. Where these are missing, archaeological data is difficult to find and access, and data re-use in collaborative research across institutional and national as well as disciplinary boundaries is limited.

Common and integrated e-infrastructure basically means accessible digital archives with structured, interoperable and re-usable data resources at the bottom and discovery, access and other services on top. ARIADNE provides e-infrastructure services which allow interoperability of existing and newly built archaeological data archives and, based on this interoperability, cross-archive search, access and re-use of available data.

ARIADNE does not replace any of the underlying infrastructures (e.g. institutional repositories and community-level data archives), but provides integrating functionality and services on top of them. Thereby ARIADNE helps making currently isolated archaeological data more accessible and useable for the research community as well as other groups such as heritage management agencies and citizens. The main intended user group of the ARIADNE e-infrastructure and services is the archaeological research community, including of course research data curators and providers.

Identifying relevant services

Much care has been devoted to identifying the services end-users such as researchers and data managers expect from the ARIADNE data infrastructure and portal. This identification has been carried out through an online survey on user needs and a survey and assessment of existing data portals (ARIADNE 2014 and 2015).

In brief, the online survey (with over 600 respondents) made clear that most researchers expect from ARIADNE to provide a data portal that allows an overview of existing archaeological data resources, and to search across the resources, using novel mechanisms for data discovery and access. Much less interest exists for typical features of Web 2.0 platforms such as content tagging or rating by users. However researchers appreciate effective mechanisms which save time in identifying relevant data, e.g. data filtering based on subjects, places and cultural periods of archaeological finds.

Much further and more detailed insights for the development of the ARIADNE data portal have been acquired through a survey of existing portals and suggestions for the project’s portal given by a panel of 23 archaeologists and data managers (ARIADNE 2015a and 2015b: section 9.3). The 34 suggestions

of the survey report have been evaluated by 28 experts of 21 partners in order to focus on the most relevant portal services. Hence, the project had a solid basis for developing a data portal that serves the immediate and evolving needs of the archaeological research community.

In general, the service portfolio of the ARIADNE infrastructure and portal should meet core requirements of data overview, search and access. This includes data search based on geo-location (maps) and date-ranges/chronologies, which the evaluators appreciated most. Not appreciated were personalized services or support of expert networking and discussion on the data portal, which is being offered by platforms such as Academia.edu, ResearchGate and others.

A “front-runner” category are services that support online research work (e-research), i.e. virtual research environments (VREs) for sub-domains of archaeology. This was not an immediate concern, but may emerge in the 10 year innovation horizon. We assume that the needs and requirements of the archaeological research community will evolve towards e-research capability when more open data becomes accessible through state-of-the-art community archives.

ARIADNE e-infrastructure architecture and services

This report does not give a detailed description of the architecture and services of the ARIADNE e-infrastructure. In brief, the core services offered by ARIADNE to the community are the Registry for archaeological datasets and the Portal for cross-archive data discovery and access.

The Registry has been implemented based on the Data Catalog Vocabulary (DCAT) recommendation of the World Wide Web Consortium (W3C), adapted for describing archaeological data resources. This adaptation is the ARIADNE Catalogue Data Model - ACDM. The Registry allows archaeological data providers, large and small, to describe their resources (collections, datasets) based on a common model (the ACDM), which enables interoperability of the resources. Data search, visualization and access on the ARIADNE data portal is now possible based on thematic, geo, temporal and other information contained in the metadata of the data resources. Further integration of data can be achieved if providers map their databases to the extended CIDOC Conceptual Reference Model.

The ARIADNE portal also provides access to special web services for images and models of research objects. These services allow easy web-based publication, visualization and exploration of high-resolution images, reflection transformation images (RTI) and 3D models, including large 3D landscapes. Other, more experimental stage applications exploit Linked Data, where semantic relations among data are expressed using the CIDOC-CRM and other domain vocabularies.

A major step towards integrated archaeological data and e-research

The overall objective of the ARIADNE e-infrastructure and data portal is providing an environment for services that act as brokers between archaeological data providers and users. The creation of such an environment is a substantial step forward in the archaeological domain, in particular it provides a common platform where dispersed data resources can be uniformly described, discovered and accessed.

The ARIADNE e-infrastructure allows overcoming some of the idiosyncrasies of the underlying different infrastructures of the data providers which currently prevent the collaborative exploitation of available data. At the same time, it represents an essential step towards the more ambitious goal of providing integrated services and tools capable to support web-based research aimed at creating new knowledge (e-archaeology).

This may be achieved by future Virtual Research Environments (VREs) built on top of the ARIADNE e-infrastructure and services. It is envisaged that e-research/science will become an important topic in

the medium to long-term as more archaeological open data and useful tools for data integration and exploitation become available (see 10-year innovation horizon).

Enabling data integration beyond ARIADNE

ARIADNE aims to enable data integration and added value services beyond the ARIADNE data portal. This can allow the initiative to play a significant role in the data service environment in and beyond the different domains of archaeological, cultural heritage and other humanities research. A core element of this scenario is Linked Open Data (LOD). ARIADNE prepares such data for interlinking with external data resources and also developers of other initiatives may use it, for example, to integrate ARIADNE-mediated and other data in the LOD “cloud” and, thereby, promote cross-domain usage.

Such interlinking is already feasible for cultural heritage data based on the CIDOC-CRM and other widely used domain vocabularies (e.g. Arts & Architecture Thesaurus). Deep interlinking with other domains relevant to archaeology, e.g. earth, environmental, biological and other sciences, will require targeted efforts to integrate conceptual knowledge and vocabularies of the different domains. Such integration could be sought based on use cases with a clear added value for archaeological research communities. While this is not a priority for ARIADNE in the 5-year horizon, we assume that in the 10-year innovation horizon such integration may be attempted.

Suggested actions: Provide core and additional data services

The actions suggested for the ARIADNE e-infrastructure services in the Preliminary innovation agenda concerned two main goals: to provide the services for the archaeological research community identified in the user needs surveys, and to enable the e-infrastructure act as a node in the wider information ecosystem of archaeology and beyond. The Preliminary innovation agenda was elaborated in the years 2014/2015 and in the meantime the first goal has been achieved. Therefore only the suggested actions for the second goal are retained.

However, a summary of the recommended ARIADNE service portfolio seems appropriate. In brief, the archaeological research community expected from ARIADNE:

- *A data portal that allows an overview of available but dispersed archaeological data resources;*
- *Capability to search across different digital archives/repositories which hold such resources (i.e. data collections, databases, datasets of projects, etc.);*
- *Effective data discovery, browsing and filtering mechanisms, in particular based on geo-location (maps) and date-ranges/chronologies, but also other advanced options such as faceted search;*
- *Data access methods according to the different data access levels, data types/products and interaction modes offered by the providers.*

These requirements have been met by providing the user community with an online facility to register and describe accessible data resources based on a common model, and offering a portal with the desired set of data discovery and access functionalities. The data access levels of providers vary, so that the portal can in general enable collection-level but in many cases not yet direct access to and interaction with data items. Thus there remains quite some potential for dataset providers and ARIADNE to enable advanced access to and interaction with data on the ARIADNE portal. Such potential is addressed in the scenarios for service extensions in the ARIADNE sustainability plan (*see Section 4.6*).

In addition to the services mentioned above, the ARIADNE portal also provides access to special services for which a strong demand has been expressed by archaeological researchers (e.g. in project workshops). This concerns the offered services for easy web-based publication, visualization and exploration of visual media such as 3D models of objects, built structures and landscapes.

Not implemented in the ARIADNE portal were services for which the user surveys evidenced low interest, for example, academic/professional profiles of users or expert networking and discussion on the portal. Such services are provided by widely used portals such as Academia.edu and others. Clearly the service portfolio of the ARIADNE portal should meet core requirements of data overview, search and access. There was little scope to invest limited funds on specific services which are not appreciated, are provided by other portals, or may run ahead of the needs of broad user segments.

Suggested actions: Data integration and services beyond the ARIADNE portal

The suggested actions concern the objective to enable the ARIADNE e-infrastructure act as a node in the wider information ecosystem of archaeology and beyond. While the ARIADNE e-infrastructure is a substantial step forward in the archaeological domain, generation of added value beyond the data Registry and Portal is envisaged. This can be achieved by adopting the Linked Open Data (LOD) approach to interlink ARIADNE-mediated data with various other data resources in the so called LOD Cloud. Such resources can be cultural heritage and other humanities data as well as from other domains (e.g. environmental, social or biological sciences). Data interlinking can be done also by project external developers of applications based on Linked Data offered by ARIADNE, provided vocabularies are available which allow for such interlinking.

The suggested actions below comprise of two recommendations concerning wide interlinking of ARIADNE-mediated and other data, and two others which do not necessarily require a Linked Data approach. But they may also allow generating significant added value within and beyond the ARIADNE portal.

Deploy Linked Open Data (LOD) to integrate information within the portal and to link to external resources which follow LOD principles

Using Linked Open Data (LOD) for information integration has been among the highest valued recommendations of the portals survey. This recommendation was followed in the development of the ARIADNE services, e.g. mapping of subjects of datasets to a common thesaurus available as LOD (the Art & Architecture Thesaurus). Use of such vocabularies will also allow interlinking with external resources which follow LOD principles. Rich interlinking with relevant resources within the domains of cultural heritage and humanities as well as of other disciplines will require dedicated efforts. To expand the initial web of ARIADNE LOD, related developer communities could be invited to create link-sets (i.e. sets of links between LOD resources).

Provide interfaces that allow external applications exploit available data, metadata and conceptual knowledge

This recommendation can be followed easily by allowing external application developers access and use the ARIADNE LOD, i.e. the SPARQL endpoint of the Linked Data [RDF] triple-store or a “dump” of the whole dataset. Available and well-documented interfaces can promote experimentation and generation of novel applications. Such applications may combine data from different sources, e.g. data “mash-ups” or assembled virtual collections of research objects. In the medium to long term we may also expect Virtual Research Environments (VREs) which use the ARIADNE LOD and others to support cross-domain collaborative research.

Enable integrated access to data and publications, e.g. include metadata of document archives and publishers

ARIADNE focuses on data resources while aggregation of metadata of publications from document repositories (i.e. self-archived papers) and publishers is not considered yet. But the connection of research publications and underpinning data is of vital importance in the scientific enterprise.

Therefore the project should investigate how interlinking of and integrated access to publications and data on the ARIADNE portal might be achieved. Established data citation standards (e.g. DataCite) will be instrumental in this regard.

Help enrich community websites with information about relevant data, e.g. RSS feeds on updated and newly available datasets for research on particular subjects, regions or periods

Such services were not appreciated by several project partners. There is a concern that this would add to the proliferation of specialised websites which are often not maintained and enriched with user contributions. However, established websites of research communities might benefit from ARIADNE information, e.g. notification about relevant new data. This would also allow enlarging ARIADNE's footprint in and beyond the sector and may stimulate re-use of available data.

3.3 10-year innovation horizon

The 10-year horizon section of the innovation agenda and action plan first addresses the goal of enabling advances in knowledge and societal relevance of archaeological research based on open and innovative approaches. Second, it points out a possible shift from the current focus of enabling access to open, interoperable and re-usable data to further progress in digital, ICT-enabled archaeological research (e-archaeology). Finally, the selected innovation topics and suggested actions in the 10-year horizon are presented.

3.3.1 Towards open and innovative e-research

The current debate about the system of scholarly research and communication is all about “open” including, but not limited to, open access (publications), open data, and linked open data. As many proponents suggest, “openness” provides much potential for novel forms of research collaboration, including participation of citizens, and innovative generation and publication of new knowledge. Thus the request for “openness” (or “open science”, “science 2.0” and other labels) is closely tied to expected innovative changes in research practices which could lead to favourable outcomes as envisioned by different proponents (Fecher & Friesike 2014). These include advances in knowledge and societal relevance of science.

The ARIADNE Innovation Agenda and Action Plan are not about incremental innovations in archaeological research through small improvements of research routines and tools. The goal, in the long term, is transformative innovations, substantial changes in archaeological research practices, with a focus on digital resources and ICT-enabled research (“e-archaeology”). A breakthrough in the next few years towards wide sharing and re-use of open data would be a transformative innovation.

The impact of the request of open or FAIR data, data which are Findable, Accessible, Interoperable and Re-useable (FORCE11; Wilkinson *et al.* 2016), is already felt in the ecosystem of research. Because compliance requires efforts such as negotiation of open data mandates, implementation of appropriate digital archives, solving intricate questions of IPR & licensing, and training of researchers (e.g. data management planning). But open data is only one element of “open science”, which is about new perspectives and ways of doing research and, hopefully, generation of new knowledge to address societal, environmental and other issues.

New technologies may trigger innovation in research practices, however, technology is but one factor in innovation and not necessarily the most important. Therefore the 10-year horizon of the ARIADNE innovation agenda is not only or primarily about technologies, although some topics require addressing questions of technology and data in greater detail.

Transformative innovation concerns the whole ecosystem of archaeological research and communication, including institutional, technical and socio-cultural dimensions. As an example: if we think of novel forms of digital publication, the core challenge is not technical but institutional. Because publication is tied to the system of scientific review, recognition and reward, and what does not fit with the established system will find it difficult to advance from prototypic solutions to wide adoption by the research community.

3.3.2 Shift from data- to research-focused innovations

The 10-year horizon of the innovation agenda builds on and extends the framework of the medium-term horizon. In line with ARIADNE's focus on research e-infrastructure and services, the topics suggested for the 10-year horizon mainly concern innovations in digital, ICT-enabled archaeological research, "e-archaeology". We assume that without substantial achievements with regard to the 5-year objectives further progress in e-archaeology is unlikely. Alternatively, if the immediate goals are met sufficiently, future-oriented perspectives will be strengthened or emerge in the 10-year horizon.

These perspectives will be less dominated by the need of growing the stock of FAIR data. This will still require much attention and improvement. But instead of being blocked by seemingly unsurmountable barriers, further progress in digital archaeology can be foreseen, explored and targeted. Scenarios in the long-term horizon of 2026 can look more into potential research-focused innovations, explore emerging new capabilities of research, and suggest pathways towards innovative and potentially transformative e-archaeology.

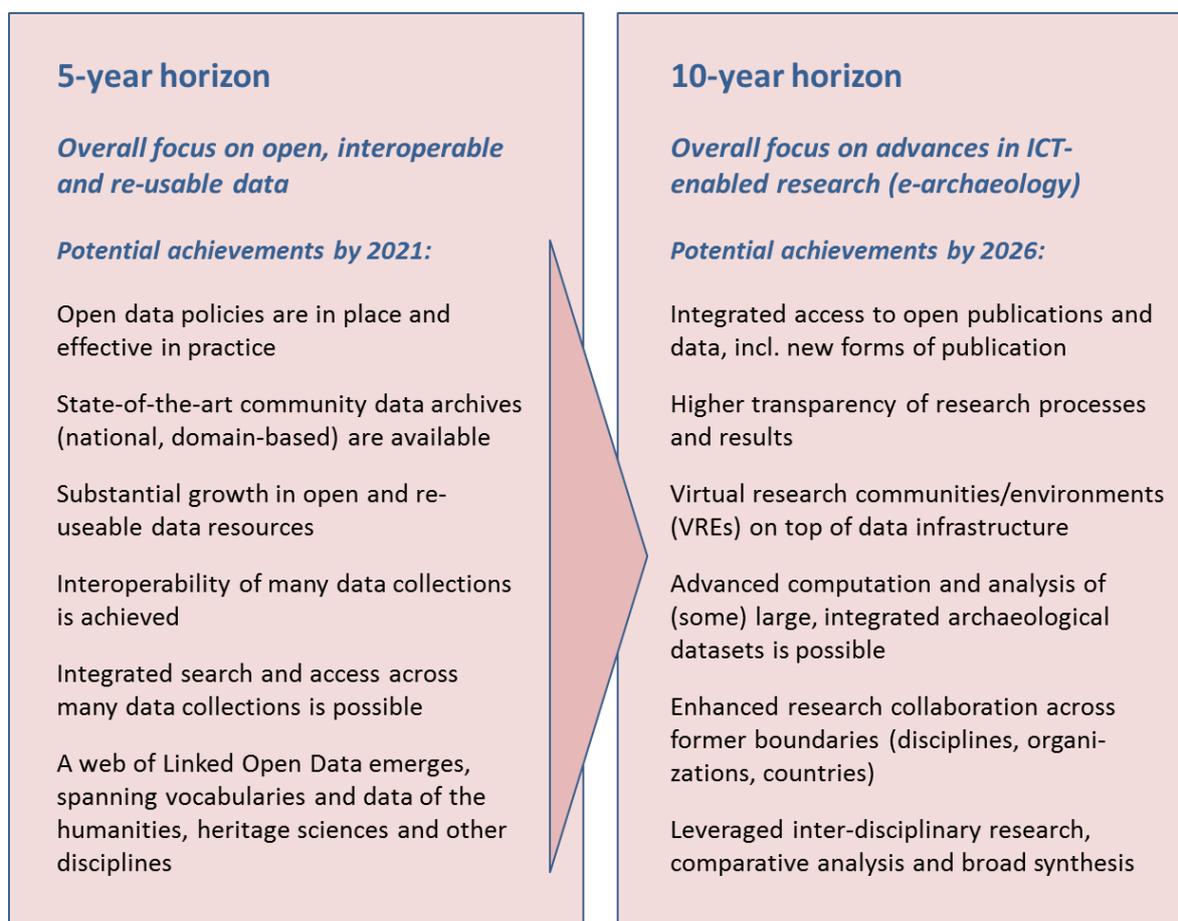


Figure 3: Towards open and transformative e-research. ARIADNE 2016.

We assume that the needs and requirements of the archaeological research community will evolve towards e-research/science capability when more open data becomes accessible and re-usable through community archives, data infrastructure and services. We do not expect, however, that such novel research practices will be driven mainly by advances concerning data and technology, changes in research collaboration and methods will be just as or even more important.

Figure 3 illustrates the shift from the current focus of enabling access to open, interoperable and re-usable data to advances in e-archaeology. It contains potential achievements which may require some explanation:

The 5-year horizon includes state-of-the-art community archives for archaeological data. At present such a data archive is missing in many European countries, hence achievement of this goal will be rather demanding. However, countries where the research community lacks a common data archive may “leapfrog” to a state-of-the-art solution by learning from acknowledged benchmarks.

Another difficult to achieve element in the 5-year horizon is the emergence of a web of Linked Open Data (LOD) that spans vocabularies and data of the humanities, heritage sciences and other disciplines. As a multi-disciplinary field of research archaeology could benefit greatly from such a LOD web as it allows discovery and retrieval of semantically related data and knowledge. But not many archaeological LOD resources have been published and interlinked as yet. In the coming years the domain of archaeological and other heritage research will have to adopt the LOD approach much more so that a rich web of LOD emerges (cf. ARIADNE 2016).

In the 10-year horizon we expect that virtual environments for archaeological research will become available, built on top of the ARIADNE and related data infrastructures and services. The archaeological research communities currently lack such environments, which should be tailored to their specific requirements. Concerning expectations of advanced computation and analysis of large integrated archaeological datasets there are some doubts. Archaeological data is difficult to integrate for such computation and “big data” mining technologies may not be applicable to many data resources.

In the 10-year horizon there will still be a need to broaden the access to various data resources required for archaeological research. As a multi-disciplinary field of research, archaeology depends on a high-level of openness of all involved disciplines and research specialties. Furthermore, required is good access to data which are usually not produced by archaeologists or in an archaeological context, airborne or satellite remote sensing and imaging data, for instance.

3.3.3 Innovation topics and suggested actions

The 10-year innovation agenda concerns a number of topics which have been identified through a scanning of recent sources (research papers, reports, project websites, etc.), with a focus on digital archaeology in its various forms. Also the expert forums of the ARIADNE Special Interest Group (SIG) on Archaeological Research Practices and Methods (2015 and 2016), and discussion of topics in ARIADNE sessions at major conferences (i.e. CAA, EAA and others) provided valuable ideas for this undertaking.

In comparison to the rather clear objectives and requirements of the next five years the 10-year horizon is a more open field. It must count on considerable achievements in the 5-year horizon, with regard to a large growth in FAIR data, for instance. Which effects these achievements will have on future e-research in archaeology are not as clear as one might think, but the drive towards innovative ways of doing archaeology will come less from research policies (as in the case of open data), rather visionary research associations, centres and groups, data scientists, tool developers and others will pave the way towards further achievements.

The ARIADNE innovation agenda for the 10-year horizon may provide some orientation, point towards interesting avenues and, thereby, help in bringing about advances in digital, ICT-enabled archaeological research. In this vein, the following actions are suggested which stakeholders could take in the next few years to promote and shape innovative and potentially transformative archaeological practices. Background and evidence for each of the suggestions is given in the preliminary version of the innovation agenda (ARIADNE 2015b).

Address the question of archaeology's societal relevance

The first point on the 10-year agenda does not address digital archaeology specifically. It reflects the increasing concern of archaeologists about the relevance of the discipline, the role archaeology can and should play in society, with regard to cultural, social and environmental challenges, for instance. This issue will hardly go away, rather become more pressing. It is not primarily about better “marketing” or involving citizens, but relevant scientific contributions to societal issues.

What would be desirable future contributions of archaeological research to tackling such challenges? What kind of research – digital, ICT-enabled or not – might enable such contributions? What is required to enable such research, for example in terms of researchers’ perspectives, knowledge, methods, technologies and data?

Concepts of digital science will be futile in archaeology (and elsewhere) if not connected to vital current debates. It is felt widely that the archaeological research community has no strong role and voice in such debates. Gaining such a role is about demonstrating relevance for societal, environmental and other concerns. This may also leverage citizens’ appreciation, understanding and support of archaeology.

Suggested actions

- *Leverage archaeology's societal role and relevance, for example through connecting archaeological research and knowledge with critical issues such as climate change, environmental sustainability, urban agglomeration, globalisation and geo-politics, regional conflicts, migration, and others.*
- *Consider contributions which could allow archaeology a stronger voice in current debates.*

Foster participatory and reflective public/community archaeology online

Many public/community archaeology projects have found it difficult to engage the communities they claimed as stakeholders in the ways described in theory, i.e. non-hierarchical, participatory or, even, “rooted” in the community. There appears to be a large gap between what is expected from involving citizens in the archaeological research process and what is actually possible in such involvement. The shift of engagement activities into the digital realm aggravates conceptual issues of public/community archaeology, for example, the (sociological) concept of community becomes illusive.

Online environments specifically built for “crowd sourcing” contributions by non-experts often present a one-way participation approach. Concerning expectations from open research data, archaeological documentation and data is not something many non-experts may easily use for own research work (“citizen science”). Mediation by archaeologists might again reproduce hierarchical and expert-directed involvement. Archaeological institutions and projects already use social media to increase visibility and disseminate information. Conceiving novel, participatory approaches based on such media will require a highly reflective usage.

Suggested actions

- *Conceive and engage in participatory approaches based on online platforms (e.g. social media) in a highly reflective way.*
- *Explore the concept and practicalities of “open research communities” that involve archaeologists and citizens in the production, dissemination and re-use of open data.*
- *Instead of seeking “roots”, focus on relevance of archaeology for current societal, environmental and other issues, at regional as well as global scale.*

Take account of the diversity of archaeological research practices and methods

Archaeology is a multi-disciplinary field of research in which researchers in various domains address their research questions with different theories, methods, data and tools, e.g. classical studies versus environmental archaeology, for instance. The diversity of “schools” and research practices in archaeology should be considered when trying to identify and conceive pathways towards innovative digital, ICT-enabled research that makes a difference. One common feature is the research lifecycle from project idea to publication of results (incl. open sharing of the generated data). Also issues of standardisation (e.g. data models, terminology) and cost-effectiveness will remain common concerns.

Suggested actions

- *Recognise that different archaeological schools of thought and research practices require different digital, ICT-based research environments and tools.*
- *Focus on phases in the lifecycle of archaeological research in which significant progress in knowledge may be achieved. In the last decades data generation has seen enormous progress, in the future other phases may require more attention.*
- *Recognise that issues of standardisation (e.g. data models) and cost-effectiveness are also relevant for future research practices.*

Target data integration for comparative and synthetic research

Open and re-usable data enable easier combination of data for comparative and synthetic, cross-disciplinary research. Enlarged and integrated datasets may allow new insights through applying models of social systems and behaviours, adaptation to climate and environmental change, for instance. In recent years archaeological research has taken aboard an arsenal of data capture methods and produced growing volumes of field survey and excavation data for documenting individual sites and areas. On the other hand, data integration for comparative and synthetic research is lagging behind ever more.

There is a need of novel approaches and tools that allow researchers bringing together and working effectively with the variety of data required for cross-domain, interdisciplinary research, in fields such as historical ecology or urban archaeology, for instance. One solution that could drive progress in large-scale data integration is setting up programmes and competence centres which support this task, for example, through facilitating collaboration between researchers with domain-specific knowledge and data science expertise.

Suggested actions

- *Foster the development of novel methods and tools that allow researchers bring together and work with the variety of data required for cross-domain, interdisciplinary research.*
- *Promote competence centres and programmes aimed to integrate data for comparative and synthetic archaeological research.*

Evolve a Web of archaeological Linked Open Data (LOD) for research

The last 10 years have seen much progress in Linked Open Data (LOD) know-how required to produce, publish and interlink LOD of archaeological and cultural heritage collections/databases. In practice, however, not many LOD datasets have been created and interlinked so far. As a multi-disciplinary field of research archaeology could benefit greatly from a web of LOD that spans vocabularies and data of the humanities, heritage sciences and other disciplines. This would allow discovery and retrieval of semantically related data and knowledge of different domains of research within and beyond archaeology.

A much wider take-up of the LOD approach for semantic interoperability in the archaeological and other domains is necessary so that a rich web of LOD can evolve. Two core requirements must be met: effective interlinking of LOD requires use of common or mapped vocabularies (thesauri, ontologies), and the LOD resources need to be curated to ensure reliable interlinking and access. A central role in the LOD scenario plays the CIDOC Conceptual Reference Model, which has been extended for scientific observations and domain-specific modelling (e.g. archaeological excavations). It is expected that mapping of archaeological data collections/ databases to the extended CIDOC-CRM will enable enhanced search as well as research-focused applications.

Suggested actions

- *Promote publication of LOD datasets (collections, databases) by more archaeological and other cultural heritage institutions, especially based on mappings to the extended CIDOC-CRM.*
- *Foster a community of LOD curators who ensure reliable availability and interlinking of LOD resources (datasets and vocabularies).*
- *Develop LOD-based applications that demonstrate advances in research capability, which may motivate a wider adoption of the LOD approach by research institutions and projects.*

Explore virtual research environments for archaeological research

The e-research scenario of virtual research environments (VREs) built on top of e-infrastructure and underlying data repositories has not yet reached the archaeological research community. The ARIADNE data infrastructure and services may inspire VRE developers to create environments for archaeological researchers. Such environments can range from loosely coupled tools and services to tightly integrated virtual workbenches for specific research communities. Relevant VREs for archaeological research should be explored, taking account of the state-of-the-art in other disciplines and particular requirements of archaeological researchers in different domains as well as in cross-domain collaboration.

Suggested actions

- *Look into VREs developed for other domains to conceive environments relevant for e-research in specific archaeological domains as well as in cross-domain collaboration.*
- *Consider cases where researchers use data mediated by ARIADNE as well as data infrastructures and services of other disciplines (e.g. geo, environmental, biological data).*

Promote new forms of media-enriched and dynamic scientific publications

The next 10 years will very likely see some advances with regard to new forms and ways archaeological researchers could publish project outcomes. However, publications are tied to the scholarly review and credit system. Without some adjustments in this system new forms of online publication may find it difficult to progress from prototypes to wide adoption by publishers and authors.

A fast take-up can be expected for moderately enhanced familiar forms of publication, e.g. embedding explorable objects such as 3D models in online papers. Advanced forms will be driven by a process-view of research rather than static documents, for example, figures which auto-update as new data becomes available. This will require deep interlinking of publications and datasets. Also the distinction between journal and repository may become increasingly blurred if digital repositories become value-adding services for scholarly publication and communication.

Archaeologists are aware that the finally “published excavation” in book form is not the optimal solution. Novel repository-based forms of publication may allow significant steps towards media/data-rich reports of investigations. We may even envision the “digital record” of an on-going excavation as a stream of data from the field and laboratories, continuously made available, analysed and discussed by subject experts. In such a setting “publications” would be snapshots of the state of knowledge at a certain time, instead of the annual excavation report and some papers by researchers in need of taking care for their academic record.

Suggested actions

- *Promote novel forms of digital publication that “work” for archaeological projects in terms of enhanced access to research outcomes as well as academic credit.*
- *Start with moderately enriched familiar ways of publication (e.g. embedding explorable digital objects in online papers), and make new approaches as easy as possible.*
- *Investigate fields of “data-driven” archaeological research and publication in which accessible datasets and executables (software, dynamic figures, etc.) play an essential role.*
- *Explore repositories as platforms for media/data-rich archaeological publications and value-adding services for scholarly communication.*

Identify e-science practices based on data infrastructure and computing facilities

E-science has become an ever broader topic which now spans all forms of research activities conducted online. The classical, still vital variant of e-science is use of advanced, distributed computing by researchers in the natural sciences and other disciplines. Research in the heritage sciences, i.e. archaeology, cultural heritage and other humanities research seldom employs such computing. Available Grid/Cloud-based Distributed Computing Infrastructures did not find much use so far by archaeologists. The ARIADNE data infrastructure and services may promote e-science activities. In future scenarios this could involve distributed computing based on a seamless flow of data to computing infrastructures and *vice versa*.

There are expectations that mining of “big data” can allow archaeologists novel insights, e.g. relevant patterns in data which suggest new research questions. However, many aggregated and integrated large archaeological datasets may not be available for quite some time. But archaeological researchers may increasingly use Cloud-based research support services for other purposes than computing, data transfer, temporary storage and access during research projects, for instance.

Suggested actions

- *Promote collaborative way-finding for e-science approaches, methods and tools relevant to archaeological researchers.*

- *Focus on e-science needs specific to archaeological research, which may differ from those of other humanities as well as the natural sciences.*
- *Look for uses of low-level Grid/Cloud based services and emerging examples of archaeological applications of “big data” mining and other methods.*

Propose grand challenges for the digital archaeology community

Experts of the archaeological informatics and computing community suggest seeking “grand challenges” which are research-directed and contribute to the development of new theories and methods (e.g. Huggett 2015). Thereby the community could play a stimulating and transformative role rather than simply support well-established research practices. Grand challenges for digital archaeology should go beyond what seems feasible in the short to medium term through applied research and engineering. The challenges would inspire the research community to push boundaries and explore new avenues of research with potential revolutionary impact, i.e. a shift in established paradigms, theories and methods of archaeological research. Grand challenges for archaeology have been identified, most of which require large-scale and integrated datasets for interdisciplinary research (Kintigh *et al.* 2014). It appears that the sought for advances in digital archaeology require bringing together domain experts (e.g. theories, methods, data) and developers (software, computing) to create novel research tools, and a strong focus on cross-domain, interdisciplinary research.

Suggested actions

- *Seek grand challenges that inspire the research community to push the boundaries of digital archaeology.*
- *Suggest challenges that promote mobilisation and integration of datasets for domain and cross-domain, interdisciplinary research.*
- *Bring together domain experts and developers to create methods and tools for such research.*

4 ARIADNE sustainability planning

4.1 Introduction and overview

This chapter presents the ARIADNE sustainability planning, its current status, because the planning and actual implementation is an ongoing activity.

To provide context, first essential related paradigms are addressed which concern open access to research results, and the role of E-infrastructures (i.e. ARIADNE) to enable access to and re-use of datasets which are deposited, but dispersed in different archives and databases. Furthermore, the goals and main activities of ARIADNE are presented in brief, and one major achievement of the project highlighted, the implementation of an e-infrastructure which enables sharing of archaeological datasets according to the FAIR principles.

Next, the main individual activities Networking/Community building, Research, and Transnational access to data services and expertise/training are discussed with regard to their potential sustainability. Then different variants concerning the maintenance and extension of the core data services (dataset registry and access portal) and pool of datasets are outlined and the basic, short-term sustainability solution presented.

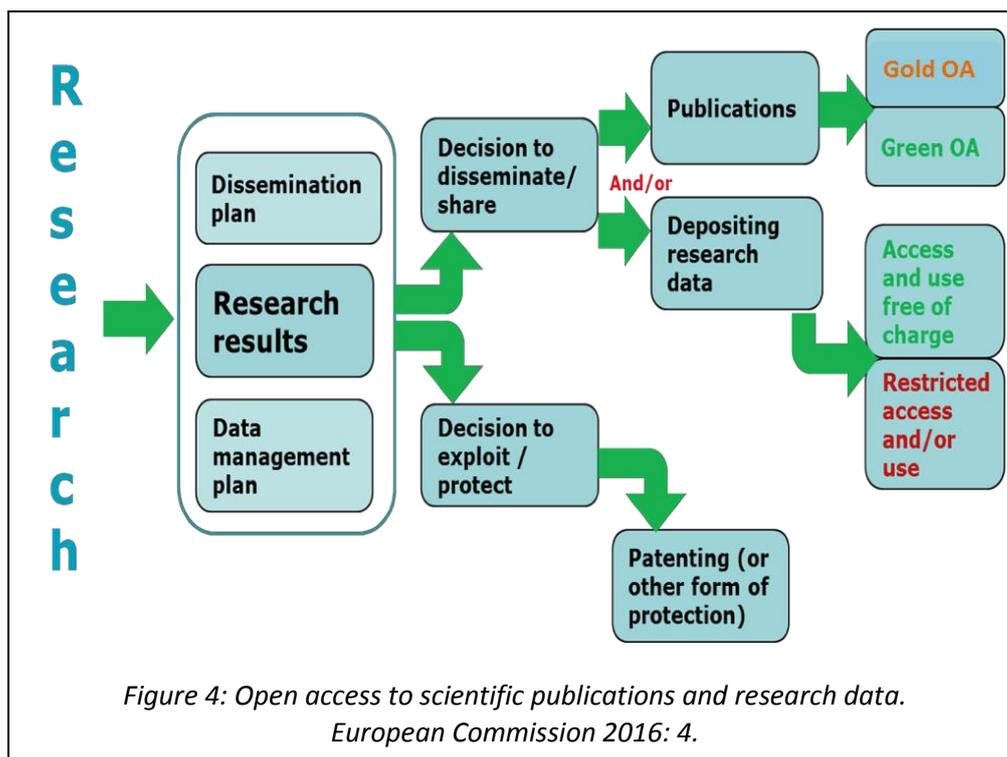
Sustainability will require extension, to keep the momentum (i.e. incorporation of new datasets of interested institutions), and to realise the full potential of the implemented system for the archaeological research community. Therefore, medium to long-term extension scenarios are presented and assessed with regard to their technical and dataset-related requirements, implied costs, and other aspects.

Finally, the current sustainability plan is presented, addressing its different elements, which include: setting up a not-for-profit ARIADNE association for further networking, training and other activities; maintenance of the project website; continued operation of the current dataset registry and access portal, and seeking opportunities and resources for their extension and long-term sustainability.

4.2 FAIR data and e-infrastructures

4.2.1 Open access to research results

The use and re-use of scientific research results is an issue that is currently at the heart of many discussions and initiatives. Openness of results is nowadays a principle which is widely accepted by the scientific community, i.e. about 600 universities, research institutes, academies, associations and others have signed by now the *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities* (2003). The European Commission is mainstreaming Open Access to scientific publications and research data as a rule to be adopted in all its funded projects. The Commission's approach to Open Access is illustrated in the following figure:



The EC Guidelines introduce data management planning, data deposit and access as part of the scientific policies on data. It is however apparent that depositing research data in an openly accessible repository is just the first step for access. Fundamental steps like providing appropriate metadata and search tools are necessary to avoid that data availability just adds more confusion and hence does not contribute to data re-use and exploitation. Data curation is necessary to ensure that data are properly stored, managed and preserved over time.

Thus, there are a number of requirements which must be fulfilled in order to guarantee that data availability is a real support to knowledge discovery and scientific progress. Such requirements are well summarized by the FAIR principles: data must be Findable, Accessible, Interoperable and Re-usable (FORCE11; Wilkinson *et al.* 2016). These principles are domain-independent and can be applied to a wide range of research outputs, but do not prescribe the way open access should be organised, i.e. how to store, share and re-use data across disciplines and borders.

While the general principles of open access to research data are common to all scientific disciplines, humanities and cultural heritage research will probably need to further specify them. Such specifications would take account of specific characteristics of their methods, most commonly the historical method, and of the ways research data are being generated, used and published, where the “raw data” often is content held (and increasingly digitised) by libraries, archives and museums.

Other data-related aspects of research in the humanities and cultural heritage, including archaeology, are generally the same as in other disciplines. The researchers share their results with others only when they are mature for final publication. Data are often considered as a means, not as a resource which could be extended and used collaboratively. This leads to the creation of a large number of relatively small datasets, the so-called “long tail of science”. They often remain isolated witnesses of work done, mostly inaccessible “dark data”, little used to support results (let alone to offer a way to check them), and almost never re-used by others because unsuitable for re-use.

There is, in sum, much work to do to foster a culture of data sharing and re-use. This culture should be based on common scientific values as well as ways data are documented and curated, for example, to enable source criticism if data created by others are considered for re-use. The work

required to overcome barriers to data sharing and re-use is addressed in the *Sections 3.2.2-4* of this report (in greater detail in the preliminary report, ARIADNE 2015: chapters 4-6).

4.2.2 E-infrastructures and the open data challenge

Research Infrastructures (RIs) are aware of the challenges of open data and are working to address them, because investment in such infrastructures, especially their e-infrastructure and services component, makes little sense without open sharing of data. Therefore they must foster and support this sharing by the research community in the domains they support. Core technical functions of e-infrastructures are to enable discovery, access to and re-use of research data held by different domain archives and databases. In addition, e-infrastructures provide services and tools to carry out research tasks, i.e. combining and asking new questions of shared data.

Particularly important in this context are the major domain RIs on the Roadmap of the European Strategy Forum on Research Infrastructures (ESFRI) as they are implemented with a perspective of long-term sustainability. DARIAH, the ESFRI RI for digital humanities, is coordinating the development of methods and systems in this field aimed to support the interconnection of datasets. Tools developed by its national nodes are shared to support the work scholars are expected to do, collaborative annotation of content, for instance. E-RIHS, the European Research Infrastructure for Heritage Science, recently included in the ESFRI Roadmap (ESFRI 2016: 52), has set the management of digital archives, their so-called DIGILAB, as one of the four pillars of the RI activity. This one has a distinctive feature: while the other pillars, the LABs, are based on activities of the network of state-of-the-art laboratories, DIGILAB will be a virtual system for accessing various databases, data repositories, digital libraries, etc., which are not necessarily managed by the laboratories.

When launching E-RIHS, the ESFRI has acknowledged the importance of ARIADNE and its liaisons with both E-RIHS and DARIAH. The ESFRI Roadmap 2016 notes that ARIADNE represents the background of E-RIHS with regard to research infrastructure integration in archaeology, more specifically, *“In the archaeological sciences the ARIADNE network developed out of the vital need to develop infrastructures for the management and integration of archaeological data at a European level. As a digital infrastructure for archaeological research ARIADNE brings together and integrates existing archaeological research data infrastructures so that researchers can use the various distributed datasets and technologies. ARIADNE has also strong links with the ESFRI DARIAH ERIC”* (cf. ESFRI 2016: 52 and 175).

As a matter of fact, the ARIADNE community is participating in the E-RIHS Preparatory Phase bringing its expertise in heritage data management and proposing its Catalogue Data Model (ACDM) and metadata standards for archaeological datasets as exemplary models for the overall domain of heritage science. The e-infrastructures cluster project PARTHENOS currently works on extending them to the whole sector of humanities to support cross-disciplinary research.

We envisage that in due course all humanities e-infrastructures will be integrated in the European Open Science Cloud (EOSC). The initiative for this all-encompassing virtual environment for research data from all disciplines has been launched by the European Commission in April 2016. In October 2016 the High-Level Expert Group on the EOSC released a report giving recommendations to the Commission on how to prepare, govern and implement the EOSC (Realising the EOSC 2016). Consequently the EOSC entered into real life, pilot actions are currently being funded to start the EOSC implementation project.

4.3 The ARIADNE Integrating Activity for archaeology

4.3.1 ARIADNE project activities

ARIADNE is the EU-funded Integrating Activity aimed at overcoming the fragmentation and isolation of archaeological datasets in Europe, fostering their interoperability and accessibility for research across institutional and national as well as disciplinary boundaries. According to the Integrating Activity scheme the project carries out a combination of three activities, Networking and community building, Research and technical development, and Trans-national access to research centres and online services, all supporting the overall goal mentioned.

Networking and community building: fosters and supports the involvement and participation of institutional stakeholders as well as individual researchers in the project activities. Cooperation agreements have been signed with major institutions across Europe (and beyond) and thousands of researchers and practitioners have been reached and involved through user needs & requirements surveys, project sessions and individual presentations at international and national sector conferences, technical and data management tutorials, participation in meetings and workshops of other projects, and so forth.

Research and technological development: concerns all activities required to enable advanced data interoperability, integration and access. This includes data models and vocabularies, e.g. the Data Catalogue Model, CIDOC-CRM extended for archaeology, database and vocabulary mapping tools for linked data. Furthermore, of course the actual building of the e-infrastructure and services, including dataset registration, ingestion and integration, and portal and other services (e.g. cross-archive data search and access, special support of visual media such as 3D models, etc.).

Transnational access (TNA): comprises of online access to data services, first provided by individual and then integrated services through the ARIADNE portal, and physical access of researchers to centres of expertise of the ARIADNE consortium. The TNA activities consist of a mixture of research and networking: they support researchers in creating their own archaeological datasets according to good practices and in view of inclusion in the ARIADNE registry and interoperability with other datasets. In a general sense, these activities are training-by-doing: face-to-face interaction with the ARIADNE research staff and joint work on data documentation projects develops skills, enlarges the community supporting ARIADNE and, indirectly, its coverage of the archaeological digital documentation.

4.3.2 ARIADNE data-related achievements

The project's Final Report on Project Impact (D2.5, January 2017) will cover all activities and present a detailed account of the results achieved. Here we highlight ARIADNE's achievements with regard to overcoming the isolation and inaccessibility of archaeological datasets dispersed in various repositories and databases of research institutions and projects across Europe. ARIADNE has created an e-infrastructure enabling online access to archaeological data which fulfils the FAIR principles:

- Archaeological data are now *findable* through the ARIADNE Portal, based on the ARIADNE Registry, which provides a comprehensive and effective system to search archaeological datasets online;
- Archaeological data are now *accessible* through the Registry, which links dataset metadata to the original content in the providers' facilities (e.g. data repositories);
- Archaeological data are now *interoperable*, in the broad sense of being described and presented in a unified way, and in the more strict sense of integration of repositories by a common

metadata schema and available multilingual thesauri and other standardization tools, such as the joint periodization table;

- Archaeological data are now *re-usable*, because the system includes a number of standardization services operating on the data and the datasets are identifiable as regards provenance, authorship and, ultimately, trustworthiness.

In short, ARIADNE is rightfully part of the above-mentioned EU vision of the European Open Science Cloud (EOSC), for which it covers the needs of a whole research community. With its work, ARIADNE is already able to contribute to the EOSC and to establish a starting point for embedding in it the archaeological data.

4.4 Sustainability of individual activities

Turning now to the ARIADNE sustainability plans, we first address each of the activities outlined above individually. The guiding question is how the activity could be sustained as such and, prospectively, as part of the portfolio of an ARIADNE association and/or continued project. The maintenance of the ARIADNE services and scenarios for extension with regard to datasets and technical infrastructure are addressed in greater detail in the next sections.

4.4.1 Research activities

In ARIADNE, pure research, i.e. research not merely aimed at the maintenance and update of the ARIADNE system, is regarded as self-sustainable. ARIADNE has produced important results such as, among others, the creation of a standard for archaeological digital documentation called CRM_{Archeo} as well as the companion specification for historic built structures, CRM_{ba}. Research and experimentation with these standards as well as other extensions of the CIDOC-CRM (Conceptual Reference Model) is continuing within the CRM Special Interest Group and involved ARIADNE partner institutions. This work will continue independently of the existence of ARIADNE, which of course currently provides substantial financial support. Other important developments concern Natural Language Processing of archaeological documents, mapping of metadata schemas and vocabularies (e.g. thesauri), and generation of Linked Data of archaeological datasets.

All these topics have by now attained independent life and are an important part of research on knowledge organization methods and tools concerning archaeology and cultural heritage in general. Therefore it can be reasonably assumed that the research activities will continue after the end of the funded period, looking for alternate funding sources or using resources internal to the involved institutions. Also the experimental models and demonstrators developed in the project, for example, concerning item-level integration of archaeological datasets, paved the way and indicated a direction to be pursued in the future.

For the maintenance of the present ARIADNE Registry, Portal and set of data services only minimal upgrades may be required to follow advancements of information technology. Additional research and technological development activities required to extend them are discussed in the scenarios section below. In conclusion, continuation of the research activities undertaken in ARIADNE is at present largely independent from the project continuation. It has influenced other projects like PARTHENOS and E-RIHS, which are extending to other domains ARIADNE's powerful approach and results. However, obtaining further funding would enable achieving a greater impact and allow developments toward innovative digital archaeology solutions such as virtual research environments as part of the ARIADNE e-infrastructure.

4.4.2 Networking and community building

ARIADNE has developed a large network of related institutions and researchers, creating a wide community of scholars and professionals interested in the same goal, the integration and interoperability of archaeological datasets in view of their use and re-use to support research and heritage management. ARIADNE has established collaborations with many institutions in Europe in charge of archaeological heritage management: among others, the Ministry of Cultural Activities, Heritage and Tourism (MIBACT) in Italy, the Directorate General for Cultural Heritage (DGPC) in Portugal, the Andalusia Historical Heritage Institute (IAPH) in Spain, the National Institute for Preventive Archaeology (Inrap) in France, the Institute of Archaeology (FSI) in Iceland, which are ARIADNE partners or associated to the project.

Furthermore ARIADNE has excellent relationships with European and international archaeological associations such as the European Association of Archaeologists (EAA), the European Archaeological Council (EAC) and the International Committee on Archaeological Heritage Management (ICAHM). For example, the European Archaeological Council strongly encourages organisations to participate in the ARIADNE initiative. The EAC comprises of heads of national services responsible under law for the management of the archaeological heritage in the Council of Europe member states. In their Amersfoort Agenda, setting the agenda for the future of archaeological heritage management in Europe, the Council emphasises *“the need to share, connect and provide access to archaeological information with the help of digital technologies. The key to this aspiration is to improve collaboration – we need to share rather than exchange. It is essential to encourage the development of European data-sharing networks and projects in the field of archaeology. The ARIADNE project is an excellent European initiative in this regard and participation in this project should be strongly encouraged”* (European Archaeological Council 2015: 21).

The stewardship of the ARIADNE community activities will need some form of coordination, among others to create and maintain interest groups within major associations and networks, to coordinate the presence at international conferences, to organize networking events, to produce or contribute to policy documents and so on. It is therefore proposed to set up a lightweight organization, formed by representatives of the ARIADNE consortium but open to the wider stakeholder community. As typical for such associations, most of the work is voluntary or, as concerns the limited secretarial work necessary for some tasks, may be provided by participating institutions with their personnel. The budget required to carry on this kind of work is therefore minimal and may come from the association fees, from special projects such as, for example, COST actions, or from the participation in other funded projects.

4.4.3 Transnational access – online

Access to data is being provided by the initial online TNA partners Archaeology Data Service in the UK, German Archaeological Institute (Arachne database) and International Association for Classical Archaeology (Fasti Online), and by other partners with established data services, e.g. the E-depot for Dutch archaeology of Data Archiving and Networked Services (Netherlands) or developed during the project such as the Archaeology Database of the Hungarian National Museum.

All online access services are integrated in the ARIADNE Portal. But access does not depend on the availability of the portal as the individual services are maintained by the partners with own resources. In a worst-case scenario, which is very unlikely to happen at least in the short to medium term, access to the archaeological datasets will continue, albeit not based on cross-archive data discovery and access. As enabling the latter is a core goal of ARIADNE it is of course envisaged to maintain the existing integration of services and datasets as well as to extend it with new ones. Scenarios and requirements for such extension are presented in [Section 4.6](#).

4.4.4 Transnational access – training

EU funding so far covers two cost categories of TNA activities: the participation of researchers and some TNA-related costs of ARIADNE partners who host participants. For the kind of TNA provided within ARIADNE (summer schools and individual research visits) the latter are very small: there are no laboratories or special equipment involved, so only personnel costs are affected. They are actually a very small fraction of staff costs, since they result as a mixture of research activities and training in a collaborative approach that has emerged as the most effective for participants.

This means that the part of the cost of ARIADNE TNA providers that concerns pure training, which is an additional cost for the research institution, is exiguous, because only for a fraction of the time spent knowledge transfer goes in one direction, from the trainer to the trainee. The larger part is either shared advancement or goes the other way round, into an improvement of the ‘trainer’ skills deriving from being exposed to new research questions and looking for their solution. Thus the time ‘donated’ for TNA training is minimal, while the institution gains a lot with regard to internal building and external recognition of expertise. The latter can lead to benefits, for example, participation in new projects, with much lower effort than with other means such as attending meetings or exchanging visits. Most institutions will understand this and will be glad to provide the human resources necessary for the task.

A further proof of the above is that during the project partner institutions involved in TNA complained that it is more expensive to produce the complicated reporting required by the European Commission to obtain TNA reimbursement than the actual sum received. This means that they would be happy, actually happier, to do it for free rather than spending administration time to receive a meagre contribution. This demonstrates that actual TNA costs of trainers are minimal, at least in the ARIADNE case. In sum, continuation of TNA is largely feasible, at least on the provider side.

The situation on the other side is a bit more difficult, because the end of the funded period will also end the EU financial support to participants for travel and accommodation. In the future, trainees will need to find their own funding to attend. Since EU support will no more be available as an incentive, the reward will need to be found in the activity: what is done in TNA will have to be so interesting that it will deserve to allocate resources (by trainees’ home institutions) and to advocate such allocation (e.g. by professional association).

This is not inconceivable: during the ARIADNE activity there have been cases of TNA trainees who could not benefit of EU support (usually because of the trans-nationality requirement) but participated in TNA activities nonetheless, getting sponsorship from other sources or from their home institution. Experience has proven that an accurate choice of the period (e.g. in low season) may substantially reduce subsistence costs, keeping them into affordable limits.

The skills acquired in TNA are a strong motivation for researchers and professionals and a potential factor for their career development. Effective data organisation is becoming increasingly important for research institutions, which have understood that the application of the FAIR data principles is a must, not least to prevent exclusion from national or international funding sources.

In conclusion, as far as participants are concerned, continuation will rely on the prestige of the ARIADNE training. This matter can be addressed by carefully granting a quality ARIADNE mark only to qualified and accredited institutions. The task may be reserved to the association mentioned above, which will also be in charge of organising the program, filtering participation and advertising the opportunities offered.

4.5 Maintenance and extension of ARIADNE services

The core services offered by ARIADNE to the community are the Registry for archaeological datasets and the Portal for cross-archive discovery of and access to items of the datasets. The ARIADNE portal also provides access to web services offered by institutes of the Italian Research Council (CNR-ISTI and CNR-ITABC) for visual media (e.g. reflectance transformation imaging, 3D models) and large terrain and landscape models. These services are maintained by the research institutes and therefore not part of the sustainability planning.

For maintenance/update and potential extension of the Registry and Portal services different activities are being considered which are described in the first section below. The next section then presents what appears as feasible with regard to basic maintenance and updates of the services, keeping them alive until new resources can be acquired, for example in a follow-up project. We exclude the possibility of freezing the Registry without updating the already registered datasets. This would mean that in short it becomes out-of-date as new data added to the datasets would not be included; the Registry would become itself an archaeological monument. Thus we consider the activities A.1 + B.1 below. Variants involving additional activities, which represent improved maintenance (e.g. inclusion of new datasets) and perspectives of future development, are presented and discussed in *Section 4.6*, including consideration of costs and how they may be covered.

4.5.1 Variants of maintenance and extension

This scheme presents variants of activities concerning the maintenance/update and potential extension of the ARIADNE services with regard to the technical implementation and integration of datasets. They concern the Registry, Portal and portfolio of services and the datasets. Technical and dataset related activities are of course intertwined but treated separately when discussing the requirements, potential and costs of different variants.

A: Data Registry and Portal

- A.1: Maintenance of the Registry and Portal, including storage, Internet access, interfaces and so on, without any further development, only minimal updates to follow technological developments.
- A.2: Development of the Portal to provide additional services, which may or may not require adaptation of the Registry (e.g. added value services with regard to data search, access, visualization and interaction).
- A.3: Provision of new related functions: e.g. data review, quality assurance, deposit and long term preservation of data, etc.

B: Datasets (registration and provision to portal services)

- B.1: Update of the current datasets (without enlarging their coverage through acquisition of datasets from new providers).
- B.2: Expansion by incorporating datasets of new institutional providers from countries already present or not yet contributing to the Registry.
- B.3: Coverage of new dataset categories and/or specific data properties (e.g. archaeological sciences, GIS, etc.). Such categories and properties will require specially tailored Registry features: all areas are indeed covered by the present Registry version, but with no specialization according to certain data-specific requirements.
- B.4: Advanced data integration: Full item-level interoperability of datasets based on rich semantics, enabling deep integration, exploration, visualization and other data-specific

operations, i.e. support of domain-specific virtual research environments according to data types/subjects and tasks.

- B.5: Incorporation of the “long tail” or research data, generated by individual and small teams of researchers, requiring new functions such as support of data deposit and curation.

Below we describe the basic sustainability solution which combines the activities A.1 and B.1. The other variants which include additional activities are presented and discussed in the next section.

4.5.2 Basic sustainability solution

The basic sustainability solution comprises of the activities A.1 and B.1:

- *A.1: Maintenance of the Portal:* This activity is only minimally technological. It requires some administration and some tiny fund allocation to cover storage and Internet access costs. Since both are very small, they can be provided *pro bono* by one of the large ARIADNE partner institutions.
- *B.1: Update of the current datasets, but no new acquisitions:* Here the data acquisition and management pipeline of the Registry needs to be considered. The pipeline comprises of the following steps:



- In the first step, ingestion, the dataset records of the providers, which follow the ARIADNE Dataset Catalogue Model (ACDM), are harvested from their repositories or, in a few cases, ingested directly by the Registry managers using information provided by contributors in various formats (XML, Excel, etc.). While the former can be completely automated using OAI-PMH compliant repositories, the latter requires some human work.
- The second step concerns data check. Experience has shown that for the updates of regular contributors most, if not all, of the work is automatic; new providers first require human assistance and intervention. The amount of work involved reduces over time and depends on the quality of data, IT resources and competence of the provider.
- The third step addresses data enrichment, for example as regards the subjects or cultural periods, currently using the Art & Architecture Thesaurus (AAT) and the PeriodO system, respectively. Again, the beginning of a contribution may need some assistance, for example to map a provider’s subject or period schemas to the standardised ARIADNE ones. For this, the mapping tool developed by the project can help substantially.
- The fourth step, data publication, is completely automatic. The (meta)data are converted to the Elasticsearch format to make them suitable for searching with the Portal engine. Proper filtering in the previous steps may prevent manual intervention in this phase, but it is impossible to guarantee that no human support is required at all.

Thus B.1, update of the current datasets, still requires all the above steps but can be carried out almost totally in an automatic way. Initial problems have already been solved so that only minor issues may emerge. Nevertheless, even such a minimal support needs to be guaranteed.

In conclusion, with the ARIADNE approach of keeping datasets with the institutional providers and storing only the metadata necessary for searching, coverage of the resources required for maintenance and small updates of the current services appear as feasible. They are well within what large research institutions like CNR may ‘donate’ as they are just a minimal fraction of their regular

technical effort. Therefore sustained operation of the current ARIADNE Registry and Portal can be easily guaranteed until resources for variants of extension of datasets and services are acquired. Indeed, the ARIADNE initiative has higher ambitions, which are presented and discussed below.

4.6 Sustainability based on medium to long-term extension scenarios

Extension of the ARIADNE Registry, Portal services and dataset pool is desirable and indeed part of the sustainability planning. The maintenance of the present services and update of the already registered datasets can be guaranteed. However, without extension the ARIADNE initiative would lose momentum and the full potential of the implemented system for the archaeological research community could not be realised. Hence, sustainability requires extension.

Therefore, scenarios have been conceived which go beyond short-term basic sustainability and assessed with regard to their technical and dataset-related requirements, implied costs and other aspects. The scenarios add to the basic sustainability solution the activity variants given in [Section 4.5](#) (their numbers, e.g. B.1-5, are included below). Like the basic solution they generally include activities of the ARIADNE Association (community networking, training, etc.) which are not addressed further below. Below we briefly present the scenarios, starting from the basic sustainability solution as bottom-line.

Scenario I: Basic sustainability

- *Technical services (A.1)*: Maintenance of the Registry and Portal, continuation of existing support services (Internet access, storage, etc.), minimal updates but no further technical development.
- *Datasets (B.1)*: No inclusion of datasets of new providers, but update of already registered datasets where this can be done automatically or with only minimal human support.
- *Required resources*: Minimal technical resources (storage, processing, etc.), certainly less than one FTE staff per year, sourced internally by the ARIADNE Association based on an agreement with one or more of the ARIADNE partners.

Scenario II: Improved sustainability

- *Technical services (A.1)*: As above.
- *Datasets (B1+B.2)*: Gradual addition also of datasets from new institutional providers, with a particular focus on providers from countries not yet present with datasets in the ARIADNE registry/portal. This will entail problems as the providers have to register the datasets based on the ACDM, carry out mapping of vocabulary, etc., hence requires considerable assistance, depending on the data quality, available skills and IT resources of the providers.
- *Required resources*: New providers are accepted based on specific arrangements and/or according to the availability of own or external funds. Supporting 10+ additional dataset providers per year will require 1-2 FTE staff per year (including for basic maintenance work).

Scenario III: Advanced sustainability

- *Technical services (A.2)*: Development of additional Portal services which provide added value for users but do not require revision of the present Registry. Such services could, for example, alert users who access data objects that an internal or external service is available to process and visualise that kind of data (e.g. a 3D model viewer), or provide links to publications which are directly or thematically related (e.g. via DataCite DOIs or subjects of datasets and papers). Also services to enrich external websites of the research community may be considered (e.g. information feeds on new datasets for particular subjects or geographic areas).

- *Datasets: Coverage of new dataset categories and/or specific data properties (B.3):* Means not only adding new datasets to the present Registry but adapting it according to data-specific requirements (e.g. requirements of geo-spatial/GIS data, archaeological sciences data). This necessitates also technical development to provide new Portal services (A.2) for data-specific handling, visualization and interaction. For example, considering GIS not just as access to archived data but taking advantage of their spatial characteristics in the Portal. This requires extending the Catalogue Data Model with specialised elements, designing appropriate new services, and incorporating a larger number of spatial/GIS data.
- *Required resources:* The mainly technical variant A.2 (not including B.3) entails variable costs depending on the type of desired new service, but will in any case require additional financial resources. Coverage of new dataset categories and/or specific data properties (B.3). Here the effort is certainly substantial, e.g. in the mentioned case of geo-spatial/GIS data maybe over 20 person-months. Therefore realisation of this scenario will need dedicated funding, for example as part of a larger project. Resource availability will circumscribe the possible extent of the scenario in terms of types and number of extensions. Fortunately, help is coming from PARTHENOS and E-RIHS, which are extending the ACDM model to support various humanities and heritage science data. Once this is completed, it will be straightforward to migrate from the ARIADNE to the extended common model.

Scenario IV: Paradigm shift – full item-level data integration

- *Technical services (A.2, extended):* Various, according to the scenario components outlined below, for example including semantic integration, search and reasoning over interlinked data, and support of virtual research environments (VREs) on top of or related to the ARIADNE data infrastructure and portal.
- *Datasets: Advanced data integration (B.4):* Full item-level integration of datasets, based on mapping of datasets to the CIDOC-CRM (incl. the recent extensions CRMarchaeo, CRMba, CRMsci and others), furthermore application of different domain/data-specific vocabularies, including for archaeological sciences dealing with materials and biological remains. This would enable various semantics-based services (e.g. advanced search and reasoning over the interlinked data), and support of domain-specific VREs according to data types/subjects and tasks. The scenario is not impossible as it has been experimentally demonstrated in ARIADNE. The pilot demonstrators investigated the requirements of a paradigm shift from data access at the collection-level (the level at most dataset providers) to item-level level discovery, access and use based on advanced semantic data integration and processing within the e-infrastructure. Full realization of the paradigm shift would dramatically impact on the way archaeological research creates and uses digital data, and introduce a real e-science approach in its methodology. It would be the prelude to the creation of VREs which incorporate other innovative services such as, for example, support for advanced data visualization, analysis and synthesis.
- *Required resources:* The scenario requires a massive allocation of resources on dedicated projects, including substantial research and development tasks. It goes much beyond ensuring ARIADNE's sustainability in the medium to long term, even if a funded continuation should take place. However, it could be embedded in the European Open Science Cloud vision and the related services which are just starting to be designed as pilots.

Scenario V: Paradigm shift – “long tail” research data

- *Technical and other functions/services (A.3):* Various, especially with regard to data quality management and data preservation/curation functions.
- *Datasets: Incorporation of the “long tail” of research data (B.5):* This shifts the focus of ARIADNE, which currently is on large datasets of institutions, to numerous sets of small data. In addition to institutional datasets, many sets from individual and small groups of researchers would be incorporated. The shift necessitates a capillary organisation of data management able to meet requirements such as data quality review/assurance, data enrichment and long-term preservation, so far covered by the institutional providers. The scenario requires extending the ARIADNE system with new curatorial and technical functions/services. Particularly important would be the long-term preservation function as appropriate archives for archaeological data are missing in many European countries (and elsewhere). Use of general-purpose repositories such as Figshare and others would be inadequate because of poor metadata and vocabulary support for archaeological data as well as other reasons (e.g. commercial background, no guarantee of long-term preservation).
- *Required resources:* Taking care for numerous small datasets generated by individual and small groups of researchers (i.e. establishing a central ARIADNE Archive) goes much beyond the current ARIADNE sustainability planning. It would necessitate a very large and sustained investment and, therefore, needs to be considered as a desirable but independent development.

Figure 5 (below) gives an overview of the described and assessed scenarios; each scenario includes all activities of the previous ones. The scenarios present the medium to long term ambitions of the ARIADNE initiative concerning extensions of the e-infrastructure and data services. Sustainability of the initiative requires extensions to keep the achieved momentum, e.g. interest of many institutions to contribute datasets, and to realise the full potential of the implemented system for the archaeological research community.

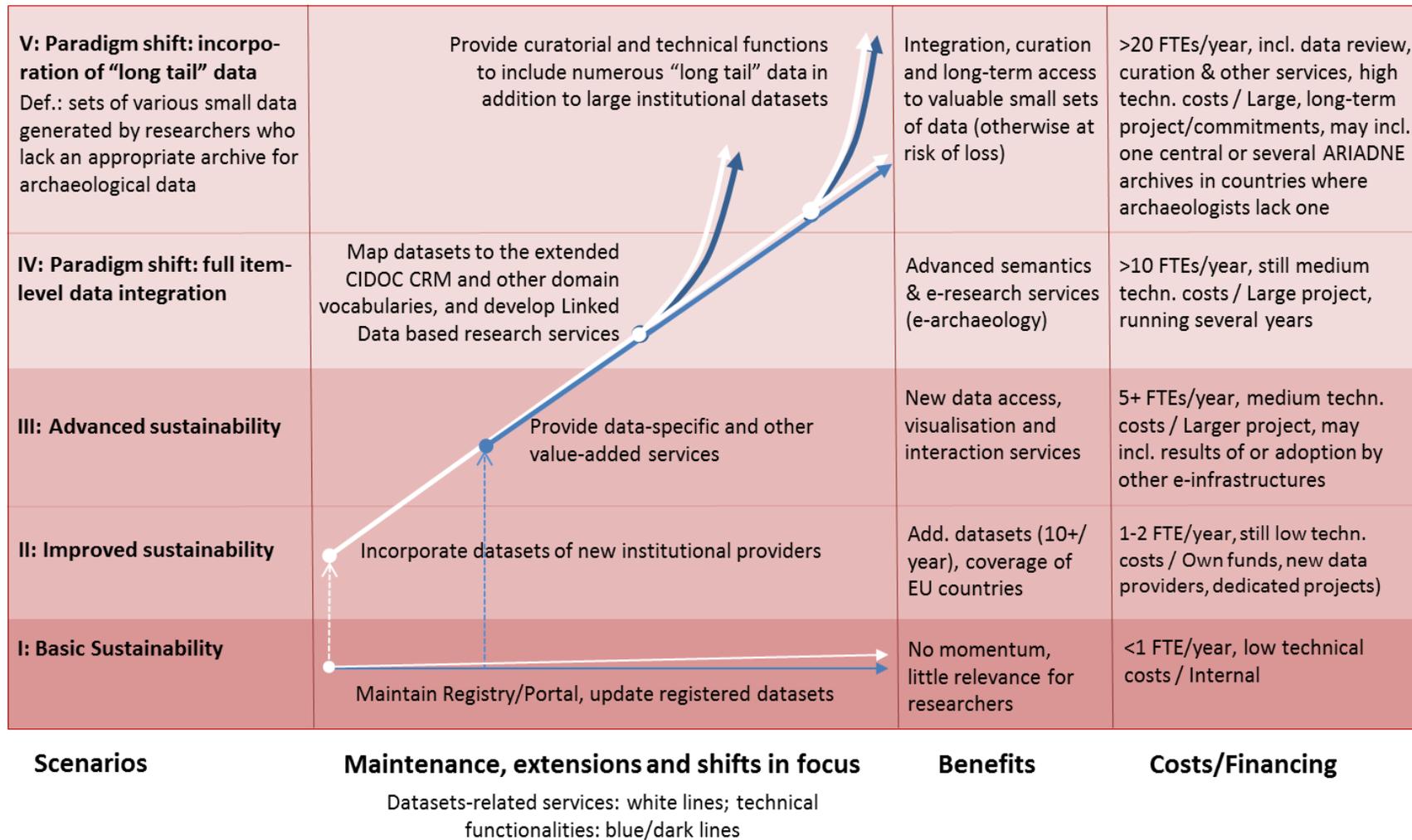


Figure 5: Overview of the ARIADNE sustainability and extension scenarios. ARIADNE 2016.

4.7 Implementing sustainability

At present, the ARIADNE plan to implement sustainability is organized as follows:

Set up an association

The association will be based on a not-for-profit scheme and, mainly for practical reason, will be established in Italy. The primary focus of the association will be to carry on the networking activities, but it will also coordinate and support other activities (see below). The founding principles will be protected by statute and it is in general expected that ARIADNE partners will have one member each. But the association members will be individuals, not institutions, as it may be difficult for institutions, especially large public ones, to participate in such an organization. The association will be open for accession of other members who fulfil certain professional criteria. The small resources required to manage the association (secretariat) will be provided by partners, at present by the coordinator. Individual activities will be supported by association members “in-kind” and acquired external funding.

Association activities

Networking: Will be the main activity of the association. This includes coordinating the ARIADNE presence at international conferences, organising network events, producing or contributing to policy documents. An important component of the networking activity will be seeking opportunities to extend the ARIADNE Registry/Portal with datasets of new providers. Also data model and technical extension are considered which, however, will require substantial, project-based funding. In such cases the Association would carry out support activities as a partner associated to a large research institution.

TNA training: The association will coordinate this activity, to be voluntarily provided by ARIADNE partners at their cost and without charging participants for the training. Support to trainees will be self-provided or availing of external funding. Some partners have already committed to continue the training as long as there is interest in the community.

Joint Research: The association will also support the research liaisons established during ARIADNE, while the research as such will continue on a self-sustained basis (i.e. not co-funded by the association).

Website

The coordinator PIN has already agreed to maintain the ARIADNE website (www.ariadne-infrastructure.eu) for a period of at least 5 years.

ARIADNE Registry and Portal

An agreement is being established with one of ARIADNE’s largest partners who will host and maintain the Registry and Portal. CNR has already declared to be available for this, and would provide the (limited) necessary human and technical resources free of charge to the ARIADNE Association. This will guarantee the Registry/Portal functionality and update for at least 5 years. Maintenance of the Registry and Portal, but lack of extension with datasets of new providers and for other data types, would restrict the usefulness of the Registry and Portal for researchers. Therefore different solutions and resources will be explored to also upgrade the ARIADNE system (i.e. the scenarios II and III described above), preventing its slow obsolescence. For example, many institutions have expressed their interest to use the Registry and Portal for publishing datasets.

Long-term sustainability

Long-term sustainability including further extension and paradigm shift (e.g. scenario IV), could be achieved based on a series of funded projects or dedicated national or international funding commitments. Research Infrastructures (RIs) on the ESFRI Roadmap operate on the basis of national commitments to enable the implementation and permanent operation of the RIs. Therefore an option for long-term sustainability of the ARIADNE e-infrastructure and services could be that a permanent RI of the humanities and cultural heritage sector takes up the system and embeds its components into their own data strategy.

DARIAH at present is not likely to offer any resources, although this possibility is being examined. E-RIHS is planning to develop a similar system and design it by 2020. Therefore it has already been agreed with E-RIHS to incorporate the ARIADNE data in the future E-RIHS Registry, if no other, better solution is found by then. In any case, to keep this possibility into account, the two systems will be developed harmoniously (e.g. E-RIHS intends to extend the data registration model developed in ARIADNE).

Other solutions for long-term sustainability may appear in the future, for example the inclusion of archaeology as one of the priority themes in the ESFRI roadmap and the upgrade of ARIADNE into an ESFRI project; but at present there are no indications supporting any expectation of this kind. However, the ARIADNE initiative has at least a 5-year period to explore various solutions for continuation and extension also of the e-infrastructure, services and pool of datasets.

5 References

Sources referenced in this report and a selection of other relevant publications and online resources.

- ALLEA - European Federation of Academies of Sciences and Humanities, E-Humanities Working Group (2015): Going Digital: Creating Change in the Humanities. ALL European Academies (ALLEA), Berlin, http://www.allea.org/wp-content/uploads/2015/07/Going-Digital_digital-version.pdf
- Anderson S., Blanke T. & Dunn S. (2010): Methodological commons: arts and humanities e-Science fundamentals, in: Watson P. *et al.* (eds.): e-Science: past, present and future, I. Phil. Trans. R. Soc. A, 368: 3779-96, <http://rsta.royalsocietypublishing.org/content/368/1925/3779.full.pdf>
- ARACHNE / IDAI.objects (central objects-database of the German Archaeological Institute, ARIADNE partner), <http://arachne.uni-koeln.de>
- Archaeology Data Service (ARIADNE partner), <http://archaeologydataservice.ac.uk>
- Archaeology Data Service (UK) and Digital Antiquity (USA): Guides to Good Practice, <http://guides.archaeologydataservice.ac.uk>
- Archéovision (Centre National de la Recherche Scientifique), <http://archeovision.cnrs.fr>
- ARCHES - Archaeological Resources in Cultural Heritage: The Standard and Guide to Best Practice in Archaeological Archiving in Europe, <http://archaeologydataservice.ac.uk/arches/>
- ARIADNE - Data portal, <http://portal.ariadne-infrastructure.eu>
- ARIADNE - Dataset Catalogue Model (ACDM), <http://support.ariadne-infrastructure.eu>
- ARIADNE - Dataset Registry, <http://registry.ariadne-infrastructure.eu>
- ARIADNE - Landscape Services (CNR-ITABC), <http://landscape.ariadne-infrastructure.eu>
- ARIADNE - Reference Model, <http://www.ariadne-infrastructure.eu/Resources/Ariadne-Reference-Model>
- ARIADNE - Visual Media Services (CNR-ISTI), <http://visual.ariadne-infrastructure.eu>
- ARIADNE (2014): First Report on Users' Needs. Project deliverable D2.1, April 2014, <http://www.ariadne-infrastructure.eu/Resources>
- ARIADNE (2015a): Second Report on Users' Needs. Project deliverable D2.2, February 2015, <http://www.ariadne-infrastructure.eu/Resources>
- ARIADNE (2015b): Preliminary Innovation Agenda and Action Plan. Project deliverable D2.3, November 2015, <http://www.ariadne-infrastructure.eu/Resources>
- ARIADNE (2016): Report on the ARIADNE Linked Data Cloud. Project deliverable 15.2, December 2016, <http://www.ariadne-infrastructure.eu/Resources>
- ARIADNE: Special Interest Group (SIG) Archaeological Research Practices and Methods, <http://www.ariadne-infrastructure.eu/Community/Special-Interest-Groups/Archaeological-Research-Practices-and-Methods>
- Art & Architecture Thesaurus (AAT), <http://www.getty.edu/research/tools/vocabularies>
- Babeu, Alison (2011): Rome Wasn't Digitized in a Day: Building a Cyberinfrastructure for Digital Classicists. Washington, DC: Council on Library and Information Resources, <http://www.clir.org/pubs/reports/pub150/pub150.pdf>

- Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003), <http://openaccess.mpg.de/Berlin-Declaration>
- Borgman, Christine L. (2010): Research Data: Who will share what, with whom, when, and why? Fifth China – North America Library Conference 2010, Beijing, 8-12 September 2010, <http://works.bepress.com/borgman/238>
- CENDARI - Collaborative European Digital Archive Infrastructure (EU, FP7, 2/2012-1/2016), <http://www.cendari.eu>
- CIDOC CRM (Conceptual Reference Model), <http://www.cidoc-crm.org>
- CINES - Centre Informatique Nationale de l'Enseignement Superior, <https://www.cines.fr>
- CLARIN - Common Language Resources and Technology Initiative (ERIC), <http://www.clarin.eu>
- COAR - Confederation of Open Access Repositories, <http://www.coar-repositories.org>
- CODATA - ICSTI Task Group on Data Citation Standards and Practices (2013): Out of cite, out of mind: The current state of practice, policy, and technology for the citation of data. In: Data Science Journal 12, 13 September 2013, https://www.jstage.jst.go.jp/article/dsj/12/0/12_OSOM13-043/article
- Costa S., Beck A., Bevan A. & Ogden J. (2012): Defining and advocating Open Data in archaeology, pp. 449-456, in: Proceedings of CAA 2012 Southampton, Amsterdam University Press, <http://dare.uva.nl/cgi/arno/show.cgi?fid=516092>
- Costas R., Meijer I., Zahedi Z. & Wouters P. (2013): The Value of Research Data. Metrics for datasets from a cultural and technical point of view. Center for Science and Technology Studies, Leiden University, <http://www.knowledge-exchange.info/Default.aspx?ID=586>
- Costs4C - Collaboration to Clarify the Costs of Curation (EU, FP7-ICT, 2/2013-1/2015), <http://www.4cproject.eu>; see also: <http://www.curationexchange.org>
- Creative Commons (CC) licenses, <https://creativecommons.org/licenses/?lang=en>
- DANS - Data Archiving and Networked Services, Netherlands (ARIADNE partner), <http://dans.knaw.nl/en>
- DANS - Digital Collaboratory for Cultural Dendrochronology (DCCD), <http://dendro.dans.knaw.nl>
- DANS - E-Depot for Dutch Archaeology (EDNA), <http://www.edna.nl>
- DARIAH - Digital Research Infrastructure for the Arts and Humanities (ERIC), <https://www.dariah.eu>
- DARIAH - Humanities at Scale/Has-DARIAH (EU, H2020, 9/2015-8/2017). <http://dariah.eu/about/humanities-at-scale.html>
- DARIAH / Buddenbohm S., Cretin N., Dijk E. *et al.* (2016): State of the art report on open access publishing of research data in the humanities. Has-DARIAH project deliverable 7.1, 12 August 2016, <https://halshs.archives-ouvertes.fr/halshs-01357208/document>
- Data Seal of Approval, <http://www.datasealofapproval.org>
- DataCite, <http://www.datacite.org>
- EDISON - Education for Data Intensive Science to Open New science frontiers (H2020 project, 8/2015-7/2017), <http://edison-project.eu>
- Edwards P.N., Jackson S.J., Bowker G.C. & Knobel C.P. (2007) : Understanding infrastructure: Dynamics, tensions, and design. Workshop report, NSF Office of Cyberinfrastructure, http://cohesion.rice.edu/Conferences/Hewlett/emplibary/UI_Final_Report.pdf

- EGI.eu - European Grid Infrastructure Foundation (coordinates a pan-European platform of distributed computing infrastructures), <http://www.egi.eu>
- e-InfraNet (2013): “Open” as the default *modus operandi* for research and higher education. Prepared by L. van der Vaart *et al.* e-InfraNet (ERA-NET) project, <https://www.surf.nl/binaries/content/assets/surf/en/knowledgebase/2013/e-InfraNet-Open-as-the-Default-Modus-Operandi-for-Research-and-Higher-Education.pdf>
- e-IRG - e-Infrastructures Reflection Group (2013): White Paper 2013. The Hague, The Netherlands, <http://e-irg.eu/white-papers>
- Elliott T., Heath S. & Muccigrosso J. (eds., 2014): Current Practice in Linked Open Data for the Ancient World. Institute for the Study of the Ancient World, New York University. ISAW Papers 7, <http://dlib.nyu.edu/awdl/isaw/isaw-papers/7/>
- ENUMERATE (2014): Survey Report on Digitisation in European Cultural Heritage Institutions 2014. Prepared by N. Stroeker and R. Vogels on behalf of the ENUMERATE Thematic Network, <http://www.enumerate.eu/fileadmin/ENUMERATE/documents/ENUMERATE-Digitisation-Survey-2014.pdf>
- EPIRIA (2014): Evaluation of pertinence and impact of the EU research infrastructure activity in FP7. A study for the European Commission by the Technopolis Group and Empirica, http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=publications
- eResearch2020 (2010): The Role of e-Infrastructures in the Creation of Global Virtual Research Communities. Empirica *et al.* Final Report, Bonn and Brussels, February 2010, <http://www.eresearch2020.eu/eResearch2020%20Final%20Report.pdf>
- E-RIHS - European Research Infrastructure for Heritage Science (since 2016 on the ESFRI Roadmap), <https://www.e-rihs.eu>
- ESFRI - European Strategy Forum on Research Infrastructures (2006): European Roadmap for Research Infrastructures. Report 2006. Luxembourg, http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri-roadmap
- ESFRI - European Strategy Forum on Research Infrastructures (2016): Strategy Report on Research Infrastructures: Roadmap 2016. Brussels, http://www.esfri.eu/sites/default/files/20160309_ROADMAP_browsable.pdf
- European Archaeological Council (2015): Setting the Agenda – Giving New Meaning to the European Archaeological Heritage. EAC Occasional Paper no. 10, Budapest: Archaeolingua, http://european-archaeological-council.org/files/amersfoort_agenda_english.pdf
- European Association of Archaeologists (EAA), <https://www.e-a-a.org>
- European Commission, DG CONNECT (2016): Cultural heritage. Digitisation, online accessibility and digital preservation. Report on the Implementation of Commission Recommendation 2011/711/EU, progress report 2013-2015, June 2016, http://ec.europa.eu/information_society/newsroom/image/document/2016-43/2013-2015_progress_report_18528.pdf
- European Commission, DG Research & Innovation (2015): Access to and Preservation of Scientific Information in Europe. Report on the implementation of Commission Recommendation COM (2012) 4890 final. Brussels, http://ec.europa.eu/research/openscience/pdf/openaccess/npr_report.pdf#view=fit&pagemode=none
- European Commission, DG Research & Innovation (2016-): European Open Science Cloud, <http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>

- European Commission, DG Research & Innovation (2016): Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020. Version 3.1, 25 August 2016, https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf
- European Science Foundation (2011): Research Infrastructures in the Digital Humanities (Science Policy Briefing 42), Strasbourg, http://www.esf.org/fileadmin/Public_documents/Publications/spb42_RI_DigitalHumanities.pdf
- EuroRis-Net Observatory of Research Infrastructures' Network, <http://observatory.euroris-net.eu>
- Fasti Online (International Association of Classical Archaeology – AIAC, ARIADNE partner), <http://www.fastionline.org>
- Fecher B. & Friesike S. (2014): Open Science: One Term, Five Schools of Thought. In: Bartling S. & Friesike S. (eds.): Opening Science. The Evolving Guide on How the Web is Changing Research, Collaboration and Scholarly Publishing, <http://book.openingscience.org>
- Fecher B., Friesike S. & Hebing M. (2015): What Drives Academic Data Sharing? In: PLOS ONE, 10(2), e0118053, <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0118053>
- FORCE11 - Future of Research Communication and e-Scholarship (n.y.): Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing (FAIR principles), version b1.0, <https://www.force11.org/fairprinciples>
- FOSTER - Facilitate Open Science Training for European Research (EU, FP7, 2/2014-1/2016), <http://www.fosteropenscience.eu>
- G8 Open Data Charter (2013) – Government of the United Kingdom, The Cabinet Office (2013): Policy paper: G8 Open Data Charter and Technical Annex, 18 June 2013, <https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex>
- Gattiglia, Gabriele (2015): Think big about data: Archaeology and the Big Data challenge. In: Archäologische Informationen, Band 38: Fokus: Open Access & Open Data, early view: http://www.dguf.de/fileadmin/AI/ArchInf-EV_Gattiglia.pdf
- Goodman A., Pepe A., Blocker A.W. *et al.* (2014): Ten Simple Rules for the Care and Feeding of Scientific Data. In: PLOS Computational Biology, 10(4): e1003542, <http://dx.doi.org/10.1371/journal.pcbi.1003542>
- GRDI2020 - Global Research Data Infrastructures (2010): Towards a 10-year vision for global research data infrastructures. GRDI2020 Consortium, <http://www.grdi2020.eu/Repository/FileScaricati/9a85ca56-c548-47e4-8b0e-86c3534ad21d.pdf>
- Harley, Diane *et al.* (2010): Assessing the Future Landscape of Scholarly Communication: An Exploration of Faculty Values and Needs in Seven Disciplines. – Archaeology Case Study. University of California Berkeley, <http://escholarship.org/uc/item/15x7385g#page-37>
- Huggett, Jeremy (2012): Disciplinary Issues: Challenging the Research and Practice of Computer Applications in Archaeology, pp. 13-24, in: Archaeology in the Digital Era. Papers from the CAA 2012, Southampton, 26-29 March 2012. Amsterdam University Press, <http://dare.uva.nl/cgi/arno/show.cgi?fid=516092>
- Huggett, Jeremy (2015): Challenging Digital Archaeology. In: Open Archaeology, 1(1): 79-85, <http://www.degruyter.com/view/j/opar.2014.1.issue-1/opar-2015-0003/opar-2015-0003.xml?format=INT>
- IANUS - Research Data Centre for Archaeology and Ancient Studies, Germany (coordinated by ARIADNE partner German Archaeological Institute), <http://www.ianus-fdz.de>

- ICAHM - International Committee on Archaeological Heritage Management, <http://icahm.icomos.org>
- IPERION CH - Integrated Platform for the European Research Infrastructure on Cultural Heritage (H2020, 2015-2019), <http://www.iperionch.eu>
- JPI-CH - Joint Programming Initiative on Cultural Heritage and Global Change (2014): Strategic Research Agenda. Rome: MIBACT, <http://www.jpi-culturalheritage.eu/2014/02/strategic-research-agenda-sra/>
- Kansa E., Witcher-Kansa S. & Watrall E. (eds., 2011): Archaeology 2.0: New Approaches to Communication and Collaboration. Cotsen Institute of Archaeology, UC Los Angeles, <http://escholarship.org/uc/item/1r6137tb>
- Kansa, Eric (2012): Openness and archaeology's information ecosystem. In: World Archaeology, 44(4): 498–520; preprint, <http://escholarship.org/uc/item/4bt04063>
- Kintigh K.W., Altschul J.H., Beaudry M.C. *et al.* (2014): Grand challenges for archaeology. In: American Antiquity, 79(1): 5-24, <http://www2.lawrence.edu/fast/PEREGRIP/Publications/2014AmAntiqGrandChallenges.pdf>
- Kvalheim V. & Kvamme T. (2014): Policies for Sharing Research Data in Social Sciences and Humanities. A survey about research funders' data policies. International Federation of Data Organisations (IFDO), <http://www.ada.edu.au/documents/ifdo-report-on-policies-for-data-sharing>
- League of European Research Universities - LERU (2013): Roadmap for Research Data. Advice Paper, No. 14, December 2013, http://www.uzh.ch/research/LERU_Roadmap_for_Research_data.pdf
- McKiernan E.C., Bourne P.E., Brown C.T., Buck S. *et al.* (2016): How open science helps researchers succeed. In: eLife 2016;5:e16800, <http://dx.doi.org/10.7554/eLife.16800>
- MERIL - Mapping of the European Research Infrastructure Landscape (inventory of research infrastructures), <http://portal.meril.eu>
- OAI-PMH: Open Archives Initiative - Protocol for Metadata Harvesting, <http://www.openarchives.org/pmh/>
- OECD (2004): Declaration on Access to Research Data from Public Funding. Paris: OECD, <http://acts.oecd.org/Instruments/ShowInstrumentView.aspx?InstrumentID=157&Lang=en&Book=FALSE>; Principles and Guidelines (2007), <http://www.oecd.org/dataoecd/9/61/38500813.pdf>
- OECD (2015): Making Open Science a Reality. OECD Working Party on Innovation and Technology Policy & OECD Secretariat, https://www.innovationpolicyplatform.org/sites/default/files/DSTI-STP-TIP%282014%299-REV2_0_0_0.pdf
- Open Data Commons (ODC) licenses, <http://opendatacommons.org/licenses/>
- OpenAIRE (since 2009, OpenAIRE2020, 1/2015-7/2018), assists the EU open access / open data policy) <https://www.openaire.eu>
- OpenDOAR - Directory of Open Access Repositories, <http://www.opendoar.org>
- Paludan-Müller, Carsten (2015): The European Association of Archaeologists. Some thoughts for the future of our Association. In: The European Archaeologist, Issue 43: 5-9, <http://e-a-a.org/TEA/TEA43.pdf>
- PARTHENOS - Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies (EU, H2020, 5/2015-4/2019), <http://www.parthenos-project.eu>
- PeriodO - Periods, Organized, <http://perio.do>

- re3data - Registry of Research Data Repositories (indexes 1500 data repositories worldwide), <http://www.re3data.org>
- Realising the European Open Science Cloud (2016): First report and recommendations of the Commission High Level Expert Group on the European Open Science Cloud. European Commission, DG Research & Innovation, Brussels, http://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf
- RECODE - Policy Recommendations for Open Access to Research Data in Europe (2015): Policy guidelines for open access and data dissemination and preservation. D5.1, February 2015, <http://recodeproject.eu/wp-content/uploads/2015/02/RECODE-D5.1-POLICY-RECOMMENDATIONS- FINAL.pdf>
- Research Data Alliance (RDA) Europe (2014): The Data Harvest: How sharing research data can yield knowledge, jobs and growth. A RDA Europe report, December 2014, <https://www.rd-alliance.org/data-harvest-report-sharing-data-knowledge-jobs-and-growth.html>
- Research Data Alliance (RDA), <https://www.rd-alliance.org>
- Research Information Network (2011a): Reinventing Research? Information Practices in the Humanities. London: RIN, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1859267
- Research Information Network (2011b): Data centres: their use, value and impact. RIN: London, <http://webarchive.nationalarchives.gov.uk/20140702233839/http://www.jisc.ac.uk/media/documents/publications/general/2011/datacentres.pdf>
- Riding the Wave (2010): How Europe can gain from the rising tide of scientific data. Final report of the European High-level Expert Group on Scientific Data. A submission to the European Commission, <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>
- ROAR - Registry of Open Access Repositories, <http://roar.eprints.org>
- Schlanger N. & Aitchison K. (eds., 2010): Archaeology and the Global Crisis. Multiple Impacts, Possible Solutions. Tervuren: Culture lab Éditions, http://www.landward.eu/2010_25Archaeology-and-the-crisis.pdf
- SPARC Europe: The Open Access Citation Advantage Service, <http://sparceurope.org/oaca/>
- Swedish National Data Service - SND (ARIADNE partner), <http://snd.gu.se>
- Taylor & Francis (2014): Open Access Survey. June 2014. Taylor & Francis / Routledge, <http://www.tandf.co.uk/journals/explore/open-access-survey-june2014.pdf>
- Tenopir C., Allard S., Douglass K., Aydinoglu A.U. *et al.* (2011): Data Sharing by Scientists: Practices and Perceptions. In: PLoS ONE, 6(6): e21101, 29 June 2011, <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0021101>
- The Royal Society (2012): Science as an Open Enterprise. Final Report, June 2012, <https://royalsociety.org/policy/projects/science-public-enterprise/Report/>
- Wilkinson M. D. *et al.* (2016): The FAIR Guiding Principles for scientific data management and stewardship. In: Scientific Data 3, 15 March 2016, <http://www.nature.com/articles/sdata201618>