D6.1 Initial Report on the Innovation Strategy and Targeted Activities

Final Version, V1.0

29 March 2020

Grant Agreement number:	823914
Project acronym:	ARIADNEplus
Project title:	Advanced Research Infrastructure for Archaeological Dataset Networking in Europe - plus
Funding Scheme:	H2020-INFRAIA-2018-1
Project co-ordinator:	Prof. Franco Niccolucci, PIN Scrl - Servizi Didattici e Scientifici per l'Università di Firenze
Tel:	+39 0574 602578
E-mail:	franco.niccolucci@gmail.com
Project website address:	www.ariadne-infrastructure.eu

ARIADNEplus is a project funded by the European Commission under the Horizon 2020 Programme, contract no. H2020-INFRAIA-2018-1-823914. The views and opinions expressed in this presentation are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission.

About this document

This document is a contractual deliverable of the ARIADNE - Advanced Research Infrastructure for Archaeological Dataset Networking in Europe – Plus project. The report presents the Initial Report on the Innovation Strategy and Targeted Activities (D6.1) that has been produced under WP6 "Fostering Innovation and Impact". The research leading to these results has received funding from the European Community's Horizon 2020 Programme (H2020-INFRAIA-2018-1) under grant agreement n° 823914. But the content of this document cannot be considered to reflect the views of the European Commission.

Partner in charge of the deliverable:	PIN Scrl - Servizi Didattici e Scientifici per l'Università di Firenze
Authors:	Guntram Geser (SRFG) Franco Niccolucci (PIN)
Contributing partners:	Elisabeth Fentress (AIAC) Sara di Giorgio (ICCU) Carlo Meghini (CNR-ISTI) Julian Richards (UoY-ADS) Douglas Tudhope (USW)
Quality review:	Holly Wright (UoY-ADS)

Table of contents

1 Executive Summary

ARIADNE/plus objectives

The overall objective of the initial ARIADNE project was to help the archaeological research and data management communities in Europe to more effectively share and reuse data resources which are dispersed and often difficult to discover and access. For this purpose the project implemented a digital infrastructure and services that enable registration, aggregation, integration, and search and retrieval of data records which describe and link to the available data(sets).

ARIADNEplus aims to take the next steps in enabling data sharing and collaborative (re)use for archaeological research across institutional and national as well as disciplinary boundaries.

The existing ARIADNE pool of data records is being extended geographically, temporally and thematically by incorporating datasets from more providers and domains of research. The records will be integrated using a Linked Data approach that enables novel ways to search and access data based on detected relations between them. Several enhanced and new research services will be provided on a Cloud-based data management and access platform. Thereby innovative and effective ways of carrying out data-based research in archaeology (e-archaeology) will be enabled.

The overall mission of ARIADNEplus is: *to integrate and effectively serve a research community that studies the past to better understand the present with the tools and the methodology of the future, in the service of European culture and society.*

The ARIADNEplus innovation strategy

The report presents the ARIADNEplus innovation strategy addressing its different dimensions. The main dimensions of the strategy are:

- *Research policies:* Alignment with the European research policies on FAIR data, Open Science practices, and the European Open Science Cloud (EOSC) initiative.
- *Data integration:* Increase of the ARIADNE data pool through incorporation of datasets from more archaeological research domains.
- *Data infrastructure:* Implementation and operation of a Cloud-based platform for data aggregation, integration, discovery, access and use across across institutional and national, as well as disciplinary boundaries.
- *Service portfolio:* Provision of enhanced and new services for digital archaeology on the Cloud-based platform.
- *Stakeholder and user base:* Extension of the stakeholder and user base in Europe and beyond, taking account of user needs regarding data, technical services and training.

Chapter 2 highlights how ARIADNEplus builds on the achievements of the initial ARIADNE project and gives an overview of the goals for the different dimensions of the innovation strategy.

Chapter 3 describes in greater detail the innovation dimensions and activities. The chapter also summarises the results and suggestions of the ARIADNEplus community needs survey regarding data sharing, access and (re)use, new services, and related training needs.

Activities and results of the WP6 task forces

The innovation strategy covers activities and contributions of all project work packages. But for some topics, which require special attention, small groups of partners in WP6 act as "task forces" that provide knowledge and guidance on how to promote innovation and impact.

The WP6 task forces address the following topics:

- e-Infrastructure synergies and joint development
- Internationalisation of the ARIADNEplus approach
- Benefits of Linked Data technologies
- Innovating services with industries

These are very different topics and, accordingly, the focus and activities of the "task forces" are also different. The focus, background, and activities and results so far are described in *Chapter 4*.

Contribution to innovation by all project activities

The ARIADNEplus innovation strategy aims to bring about significant advances of archaeological research and data management in Europe and beyond. The innovation strategy is supported by all project work packages. *Chapter 5* describes their contribution to the different dimensions of the innovation strategy, including selected activities and results so far.

From innovation to impact

Chapter 6 presents an overview of the ARIADNEplus impact indicators and describes their mapping to the impacts expected of Integrating Activities for Advanced Communities funded under the Horizon 2020 Work Programme 2018-2020 for European Research Infrastructures (including e-Infra-structures). Furthermore it gives an outlook on the further work and deliverables on project innovation and impact. These are, the first impact monitoring and evaluation report in June 2021 (full account and evaluation in December 2022), the final innovation strategy (April 2022), and the sustainability plan for the service provision beyond the funded period of the project (December 2022).

2 Overview of the innovation strategy

This chapter highlights how ARIADNEplus builds on the achievements of the initial ARIADNE project and gives an overview of the goals for the different dimensions of the innovation strategy.

2.1 Building on the achievements of the initial ARIADNE project

The initial 4-year ARIADNE project (until January 2017) was an Integrating Activity funded under the 7th Framework Programme of Research and Development (FP7) of the Europen Union. The second round, ARIADNEplus, funded from January 2019 to December 2022 under the EU Horizon 2020 Programme is also such a project. An Integrating Activity aims to integrate, within the European Research Area (ERA), the community of a field of research.

The overall objectives of the first round of the ARIADNE initiative have been to build a community of archaeological institutions in Europe interested in making their data findable and accessible through a digital research infrastructure. The infrastructure should aggregate and integrate records of data items from their repositories and databases, and provide a portal for discovering and accessing items in the distributed sources. Support for the data sharing community focused on guidance in the preparation of what nowadays is commonly called FAIR data, particularly use of common data models and vocabularies to enable interoperability for search and access.

2.1.1 Key achievements of ARIADNE

The ARIADNE Impact Report addresses all impact areas and gives a detailed account of ARIADNE's achievements (ARIADNE 2017a). The report states that the project achieved good results in all evaluation areas, and highlights that it:

- Accomplished its goal to provide a digital infrastructure and services for searching and accessing archaeological data in repositories and databases of institutions in different European countries;
- Increased interoperability of datasets based on a common model (ARIADNE Catalogue Data Model), improved vocabularies (e.g. vocabulary mapping tools), and other methods;
- Implemented a European-level data portal providing advanced search capability for 'what' (subjects), 'where' (location) and 'when' (cultural chronology / date ranges).
- Made available additional high-value services (e.g. 3D artefact and landscape services), and demonstrated advanced capability in making data more accessible and useful (e.g. fieldwork reports through metadata extraction with natural language processing methods);
- Achieved a large 'footprint' in the sector regarding the numbers of institutions and researchers that have been informed and involved, including potential providers of additional datasets.

The core of the ARIADNE project has been the building of a European-level platform where dispersed archaeological data resources can be registered, shared, discovered and accessed (Aloia et al. 2017). Such a platform did not exist before and its implementation arguably is ARIADNE's key innovation for the archaeological community in Europe (and beyond). The ARIADNE Impact Report concludes that the project not only had a strong impact, but that it could become a lasting impact, especially by exploiting the high potential for further advances provided by the data sharing and access infrastructure.

2.1.2 Realising the full potential

ARIADNE built a solid basis for the next steps in data sharing and use for research across institutional and national, as well as disciplinary boundaries. ARIADNE established a common platform for archaeological data sharing, discovery and access, and seeded it with representative datasets from repositories and databases of project partners. The digital infrastructure and services span the whole chain from data aggregation to search and access services for the integrated data.

ARIADNEplus will extend the ARIADNE data pool regarding coverage of regions, time-spans, and archaeological domains, and provide a Cloud-based platform that in addition to the data search and access portal will provide enhanced and new research-focused services.

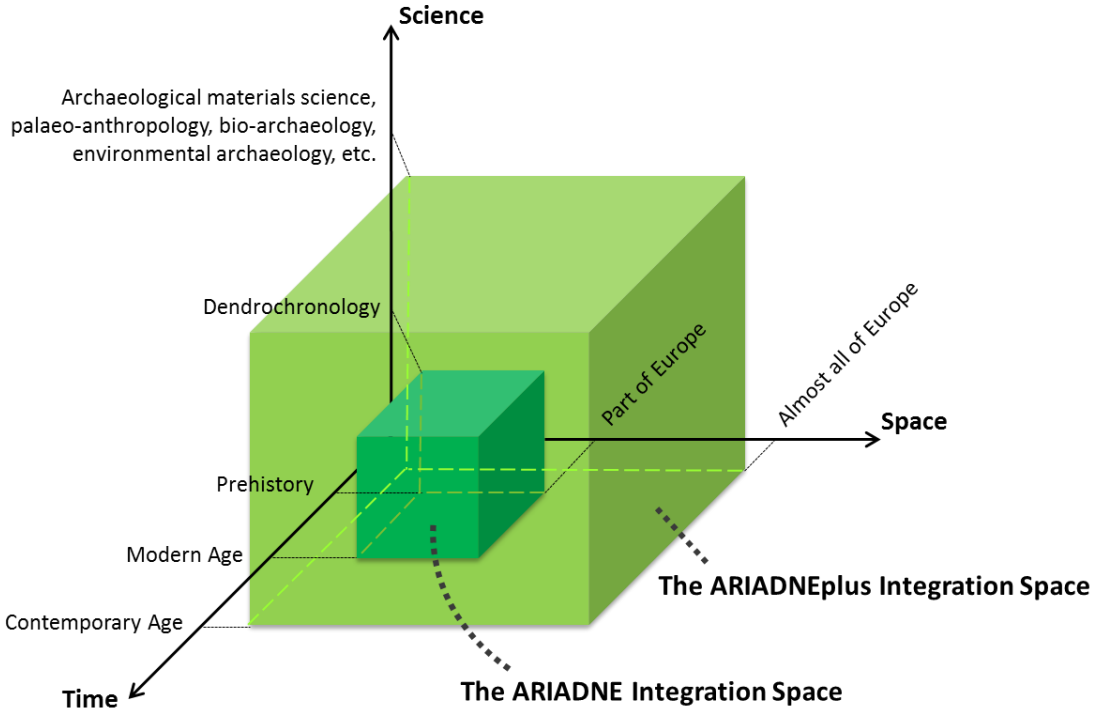

Figure 1: Extension of the ARIADNEplus integration scope in time, space and content (ARIADNEplus 2019).

Figure 1 illustrates how ARIADNEplus will progressively extend its integration scope regarding archaeological research domains, time-span and geographic coverage.

The ARIADNE consortium had 23 partners while ARIADNEplus comprises 37 partners from 23 European countries and one each from Argentina, Israel, Japan, and the United States. Thus the initiative is present now not only more strongly around Europe but also in other world regions.

ARIADNE integrated mainly data records from monument and site inventories, excavation archives, fieldwork reports and artefact databases, while few for specialist topics such as, for example, dendrochronology. ARIADNEplus will incorporate datasets from a wide range of research domains, e.g. archaeological materials science, bioarchaeology, environmental archaeology, among others. Indeed, the ARIADNE initiative aims to progressively extend its research communities and data integration scope, expand the thematic, geographic as well as time-span coverage, pushing it back to the earliest traces of human presence on Earth and forward to recent times.

2.2 Dimensions and goals of the innovation strategy

The ARIADNEplus innovation strategy aims to bring about significant advances of archaeological research and data management using different dimensions. The main dimensions are alignment with the latest European research policies, integration of datasets from more archaeological research domains, a portfolio of e-infrastructure services for digital archeology, and extension of the user base. The goals of the innovation strategy for these dimensions are summarised in the sections that follow (see *Chapter 3* for background).

2.2.1 Alignment with European research policies

FAIR data

- Develop and share expertise in the application of the FAIR data principles in the field of archaeological research and data management, including heritage management and preventive archaeology.
- Support FAIR data policies and good practice for researchers and data managers through guidelines, workshops and training (e.g. online webinars to reach many researchers).
- Provide tools for FAIR data management, e.g. a Data Management Plan template and supporting wizards; FAIR data evaluation tools.
- Work in partnership with initiatives to support national and transnational policy-making bodies in the implementation of the FAIR data principles in the field of archaeological research and data management.

FAIR repositories

- Collaborate with the SEADDA COST Action to foster the development of FAIR archaeological data repositories in countries where the research community lacks an appropriate repository.
- Help developers of new repositories plan participation in the ARIADNE data infrastructure already at an early stage.
- Enable advanced repository initiatives to benefit from available ARIADNE services, for example, by using data description and mapping services for representative initial datasets.

Open Science

- Promote Open Science practices in digital archaeology aimed to advance and make knowledge about past cultures relevant for current societal issues more accessible.
- Support Open Science practices with data infrastructures and services for collaborative archaeological research across institutional and national, as well as disciplinary boundaries.

European Open Science Cloud (EOSC)

- Develop Cloud-based ARIADNEplus data services to allow improved use of resources, economies of scale and cost-savings.
- Cooperate with other e-infrastructures for digital heritage and humanities research regarding harmonisation of data catalogues.
- Create a mapping between the ARIADNEplus catalogue model and the EOSC model for service and data resources, and register resources in the EOSC.

2.2.2 Data providers, data coverage and integration

Data providers

- Integrate data providers from more European countries into the new ARIADNEplus infrastructure.
- Give particular attention to regions in Europe where coverage was less intensive in the first ARIADNE project.
- Integrate data from non-European providers; ARIADNEplus includes one partner each in the USA, Argentina, Japan, Israel, and others are welcome.

Data coverage

Extension of the ARIADNE data pool regarding coverage of regions, archaeological domains and time-span:

- Extend the geographic coverage as stated under data providers.
- Integrate data from a wider range of archaeological research domains, e.g. palaeoanthropology, bioarchaeology, environmental archaeology, scientific analyses (e.g. material sciences), remote sensing, standing structures, among others.
- Aim to cover the full time-span of human presence on Earth, pushing the time-span back to the earliest traces and forward to recent times.

Data description, depth of data integration, and Linked Data

- Develop application profiles for data records representing different archaeological domains with domain experts, and support data providers in mapping to the ARIADNEplus CIDOC CRM based data model.
- Promote the standardised description of the metadata records with appropriate general and domain-specific vocabularies (e.g. WGS84, PeriodO, Getty AAT, scientific domain terminologies).
- Extend the interoperability of data records integrated from federated collections/databases to item-level access.
- Follow FAIR principles in building the ARIADNE Content Cloud (AC) to create a FAIR dataset of Linked Data.
- Apply Linked Data methods to exploit relations between data within the AC as well as with reliable external data resources.

2.2.3 Digital infrastructure and services

- Offer easy-to-use Cloud-based services to make it easier for researchers to acquire and manage software for research (“Software as a Service”).
- Bring IT specialists and archaeologists together to jointly develop services that support user needs regarding new research capabilities.
- Provide different data search, visualisation and access services as required for integrated data, e.g. Linked Data-based search, sites and finds distribution maps, item-level access and study with data-specific tools (e.g. for 3D models).

- Set up an ecosystem for digital archaeological research that incorporates data and services and enables use of Cloud-based Virtual Research Environments (VREs), which combine services as needed by researchers of different thematic areas.
- Demonstrate innovative uses of the new research platform and data resources in a number of pilots.

2.2.4 User base and needs

Research community involved:

- Extend the user base regarding data providers from around Europe and non-European countries (e.g. USA, Argentina, Japan, Israel) and users of services and data.
- Support users from different archaeological research domains, e.g. palaeoanthropology, bioarchaeology, environmental archaeology, scientific analyses (material sciences), remote sensing, among others.
- Take account of user needs, e.g. regarding data management and sharing, data services, and skills development.

Specific suggestions on user needs from the ARIADNEplus community survey (ARIADNEplus 2019) are presented in *Section 3.3*.

3 Detailed project innovation strategy

This chapter describes in greater detail the different dimensions and activities covered by the innovation strategy. Also included are results and suggestions from the ARIADNEplus community needs survey regarding data sharing, access and (re)use, new services, and related training needs.

3.1 Alignment with European research policies

The ARIADNEplus innovation strategy is aligned with the current European policies regarding FAIR research data, Open Science practices, and the European Open Science Cloud (EOSC) initiative.

3.1.1 FAIR data

Background

Over the last few years the FAIR data principles, published in April 2016, have been adopted by different stakeholders for the sharing and reuse of research data through data repositories and infrastructures, including the ARIADNE initiative.

The FAIR data principles require "*that all research objects should be Findable, Accessible, Inter-operable and Reusable (FAIR) both for machines and for people"* (Wilkinson et al. 2016). The 15 principles address important attributes of research data, for example, globally unique and persistent identifiers, rich metadata, use of community vocabularies, registration in a searchable resource, and release with a clear data usage license (Wilkinson et al. 2016; see also Mons et al. 2017). The European Commission has brought together an Expert Group on FAIR Data to analyse what is needed for "turning FAIR into reality" and suggest concrete actions for all stakeholders (Expert Group on FAIR Data 2018).

Compliance of research data with the FAIR data principles is seen as an important success factor for the European Open Science Cloud (EOSC), which is intended to allow researchers to store, (re)use and analyse FAIR data for research, innovation and educational purposes. Many initiatives and projects are now supporting the FAIR data agenda, e.g. FAIRsFAIR[1], GO-FAIR[2], RDA FAIR Data Maturity Model WG[3], among others. But the effort to realise and manage large volumes of FAIR data from all disciplines as envisaged for the EOSC will be substantial.

While there is a FAIR "boom" in the international research data management community no wide awareness, least of all knowledge, of the principles among researchers can be assumed. A survey in 2018 with 1,239 respondents from different disciplines found that 60% of respondents had never heard of the FAIR principles before, 25% had heard of the principles but were not familiar with them, while 15% claimed being familiar with the principles (Figshare 2018 [dataset]). The survey directors state, *"This lack of awareness is concerning as the FAIR principles are being rapidly adopted by publishers, funders and institutions worldwide but there is a crucial gap in educating researchers on what is expected of them"* (Figshare 2018 [report]: 11).

It is expected that research funders will in future make data-related costs eligible for FAIR data only. This will require a lot of investment in training of FAIR data managers of research organisations, repositories and other research infrastructure. Barend Mons, chairman of the first High Level Expert

[1] FAIRsFAIR, https://www.fairsfair.eu

[2] Go-FAIR, https://www.go-fair.org;

[3] Research Data Alliance: FAIR Data Maturity Model WG, https://www.rd-alliance.org/groups/fair-data-maturity-model-wg

Group on the EOSC, estimates that for the EOSC to be successful over the next decade 500,000 data managers would need to be trained to make research data FAIR, one data expert per 20 researchers (Mons 2016).

ARIADNEplus and the FAIR principles

The ARIADNEplus project is committed to develop and share expertise in the application of the FAIR data principles in the field of archaeological research and data management, including heritage management and preventive archaeology.

The project has a work package (WP3) comprised of six tasks dedicated to policies and good practices of FAIR data management for both archaeological researchers and data repositories. The tasks will evaluate implications of the implementation of FAIR data; provide policy support tools such as a flexible Data Management Plan template and supporting wizards; guidelines and support on repository creation, management and quality control; guidance on how to realise FAIRness of data taking account of different national regulations, IPR-related and other issues; and offer practical training materials and workshops.

From these activities significant contributions to capacity building and take-up of the FAIR principles are expected. The contributions should be as practical as possible, distinct from the broad wave of general information on the FAIR principles by ever more common on the FAIR data "bandwagon". ARIADNEplus guidance and training in FAIR data should focus on what matters for archaeological researchers and data managers specifically.

In the ARIADNEplus survey how to apply FAIR data principles in archaeology ranked at the top of suggested training offers for researchers and data managers. with 94.5% of respondents considering this as helpful or very helpful (67.3% very helpful) (ARIADNEplus 2019: 121).

FAIR data innovation activities

- Develop and share expertise in the application of the FAIR data principles in the field of archaeological research and data management, including heritage management and preventive archaeology.
- Support FAIR data policies and good practices for researchers and data managers through guidelines, workshops and training (e.g. online webinars to reach many researchers).
- Provide tools for FAIR data management, e.g. a Data Management Plan template and supporting wizards; FAIR data evaluation tools.
- Work in partnership with others (e.g. the COST Action SEADDA) to support national and transnational policy-making bodies in the implementation of the FAIR data principles in the field of archaeological research and data management (e.g. European Archaeological Association, European Archaeological Council, Society for American Archaeology).

3.1.2 FAIR repositories

Background

Sharing of FAIR research data requires trusted data repositories. This is one of the key recommend-dations of the *Turning FAIR into Reality* report which states: *"Research data should be made available by means of Trusted Digital Repositories, and where possible in those with a mission and expertise to support a specific discipline or interdisciplinary research community"* (Expert Group on FAIR Data 2018: 43).

That discipline-specific repositories will suit best the FAIR requirements is also emphasised in a recent All European Academies (ALLEA) report on FAIR data sharing in the humanities. The report recommends: *"Use disciplinary repositories where they exist, as they are more likely to be developed around domain expertise, disciplinary practices and community-based standards, which will promote the findability, accessibility, interoperability and ultimately the reuse and value of your data. The level of curation available in a repository is key to data quality and reusability"* (ALLEA 2020: 29).

But many archaeologists in European and other countries do not yet have available such a repository where they can deposit and make available their data to the research community and other users. Ideally such a repository has a national scope and is mandated by research funders for depositing data from archaeological investigations. This provides advantages in several respects, including clear orientation of all stakeholders, expertise in archiving archaeological data, cost-effectiveness of data curation and access (e.g. economies of scale), among others. From the perspective of ARIADNEplus one or only few core repositories per country from which data records can be aggregated is of course the preferred scenario (Geser 2019a; Geser 2019b: 195-196).

Benchmarks for national-level archaeological data repositories exist, for example, the Archaeology Data Service (UK) and the E-Depot for Dutch Archaeology of DANS (Netherlands), both are ARIADNEplus partners. In the United States, Digital Antiquity at the Arizona State University (also a partner in ARIADNEplus) aspires to provide a national-level repository with tDAR, The Digital Archaeological Record (McManamon et al. 2017). Digital Antiquity is a collaborative organisation that involves many distinguished scholars and institutions through their board of directors and professional advisory panel, and is being supported by the Society for American Archaeology.

ARIADNE has inspired project partners in other countries to promote the building of archaeological repositories or collections in their country (see the articles in Richards & Niccolucci 2019). In smaller countries such as Austria, Ireland and Slovenia some progress has been achieved. For example, the Digital Repository of Ireland, which has a focus on social sciences and humanities, ingested a first large collection of archaeological fieldwork documentation commissioned 2001-2016 by Transport Infrastructure Ireland (TII 2017).

In European and other countries much effort will be necessary to create more data archiving solutions so that archaeologists can safely deposit and make FAIR data available to the research community and other users. Existing document repositories within university libraries and other institutions are considered inadequate for disciplinary research data. Curation requires dedicated data repositories and curators with a background in the respective domains.

ARIADNEplus and FAIR archaeological repositories

The objective of the ARIADNE data infrastructure is to allow researchers and other users to discover and access archaeological data held and shared by repositories across Europe and beyond. Fortunately, the lack of appropriate data repositories is now being addressed by the COST Action SEADDA[4] involving ARIADNEplus partners and institutions from other countries, with representation from nearly all European countries, not only those in the EU.

The SEADDA network brings archaeologists and data management specialists together to share expertise, provide knowledge and training in matters of data archiving and access, and help archaeological communities to address problems in the most appropriate way within their own countries.

SEADDA fosters the development of archaeological data repositories in countries where the research community lacks an appropriate repository, while ARIADNEplus supports finding and accessing data

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

that is being shared through existing and new repositories. ARIADNEplus can help developers of new repositories plan participation in the research infrastructure at an early stage, so that thereafter records of FAIR data can be easily aggregated and integrated in the common pool of the ARIADNE initiative.

FAIR repository innovation activities

- Collaborate with the SEADDA COST Action to foster the development of FAIR archaeological data repositories in countries where the research community lacks an appropriate repository.
- Help developers of new repositories plan participation in the ARIADNE data infrastructure already at an early stage.
- Enable advanced repository initiatives to benefit from available ARIADNE services, for example, by using data description and mapping services for representative initial datasets.

3.1.3 Open Science

Background

The vision of Open Science is making the research process and results as transparent and accessible as possible in order to advance scientific knowledge and increase innovation and societal benefits through science and technology.

In practice, proponents have different perspectives on Open Science and emphasise some aspects more than others. Fecher & Friesike (2014) distinguish five Open Science schools of thought that centre on different goals: openly available platforms, tools and services for research; collaborative knowledge creation; making science and scientific knowledge accessible for everyone; enable participation of citizens; and new metrics for relevant impacts of scientific works.

The core but difficult to grasp element of Open Science is "openness". A report of the ERA-net project e-InfraNet discusses the concept of openness in the contexts of research, content/data, software, infrastructure, standards, and innovation. The report suggests openness as the *"default modus operandi"* for all publicly funded research and educational resources, with "open" as the preferable approach *"not as an end in itself or as an ideology"* (e-InfraNet 2013: 10).

Open Science was introduced as a priority and guiding principle of research policy at the European level by Carlos Moedas, the Commissioner for Research, Science and Innovation (2014-2019). The Commissioner adopted the concept of "open innovation" and pushed for more openness of research in the European Research Area (ERA). The title of his speech at the 2015 ERA Conference, *"Open Innovation, Open Science, Open to the World"* (Moedas 2015) has also been used for a commissioned study which describes the background and requirements for a research policy set to openness in all respects (European Commission 2016b).

Since September 2016 the Open Science Policy Platform (OSPP) expert group[5] advises the European Commission on how to develop open science policies. In April 2018 the OSPP issued recommend-dations on eight topics (OSPP 2018):

- Rewards and Incentives
- Research Indicators and Next-Generation Metrics
- Future of Scholarly Communication
- European Open Science Cloud

[5] Open Science Policy Platform (OSPP), https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-policy-platform

- FAIR Data
- Research Integrity
- Skills and Education
- Citizen Science

The recommendations are meant for research policy makers, research funders, researchers and research organisations, citizen science and public engagement organisations, research publishers and libraries, and research infrastructures, including e-infrastructures.

Most relevant for research e-infrastructures such as ARIADNEplus is that the OSPP suggests: *"Data resulting from publicly funded research must be made FAIR and citable, and be as open as possible, as closed as necessary"* (under FAIR Data), and "*a key requirement is deposition in a trusted repository that adheres to FAIR principles"* (under Future of Scholarly Communication).

The OSPP also gives recommendations for the European Open Science Cloud (EOSC), which is addressed in *Section 3.1.4.* Here we note that the OSPP suggests that in future researchers should *"be required to deposit their research outcomes in EOSC-compliant infrastructures"*, and that the EOSC *"must also support the diversity of requirements across all disciplines"*, and be compatible with other research e-infrastructures.

In European Commission research policy documents the term Open Science is now being used regularly when open access to shared research resources (publications, data, software) is addressed. The expected impact of openness is a transformation of science leading to advances in knowledge, innovation and societal benefits. Innovative digital, ICT-enabled research is understood to play a key role in this transformation. Open Science using digital methods and tools can enhance research collaboration, involvement of citizens, and transparency and relevance of better accessible research results (publications and data).

ARIADNEplus and Open Science

The concept of Open Science is of course highly relevant also for archaeological research and data access and (re)use supported by ARIADNEplus. Open Science using shared digital tools and data is expected to enable novel forms of research collaboration and more accessible research results, extending the societal relevance and reach of archaeological knowledge.

The need and challenges of open research practices and resources are being discussed in the archaeological community (e.g. Beck & Neylon 2012; Beck 2013; Costa et al. 2012; Kansa 2012; Lake 2012; Wilson & Edwards 2015). A recent publication on *Open Science in Archaeology* by Marwick et al. (2017), a large group of recognised archaeological researchers, has greatly added to the awareness of open research and data in the field.

An open data imperative is particularly strong in this field: excavation destroys primary archaeological evidence; work that is done in the public interest, and there is little commercial relevance in archaeological data. Therefore openness should become embedded in archaeological research practices as *"the default modus operandi"* (e-InfraNet 2013) so that the advantages of accessible and reusable data gain priority over the interest of the individual researcher.

Many archaeologists are not yet well prepared or equipped for the Open Science research paradigm. As the matter is complex, strong leadership regarding policies/mandates, supportive institutional measures (e.g. capacity building, training of researchers), and state-of-the-art digital infrastructures for Open Science are required.

Open Science innovation activities

- Promote Open Science practices in digital archaeology aimed to advance and make knowledge about past cultures relevant for current societal issues more accessible.
- Support Open Science practices with data infrastructures and services for collaborative archaeological research across institutional and national, as well as disciplinary boundaries.

3.1.4 European Open Science Cloud (EOSC)

Background

The European Open Science Cloud (EOSC) initiative, launched in April 2016 (European Commission 2016a), aims to provide a trusted common virtual platform for researchers to store, (re)use and analyse shared data for research, innovation and educational purposes. The overall goals of the EOSC are to remove the current lack of integration of research e-infrastructures of different disciplines, support the sharing and reuse of FAIR data resources, and promote Open Science practices (as addressed above).

A Commission Staff Working Document (2018) gives an overview of the EOSC implementation roadmap. It describes an "EOSC model" by distinguishing six lines of required implementation actions. These action lines are for *"a pan-European federation of data infrastructures built around a federating core and providing access to a wide range of publicly funded services supplied at national, regional and institutional levels, and to complementary commercial services"* (European Commission 2018a: 9). The working paper summarises the action lines in a figure that presents the general setup of the EOSC (Figure 2).

Figure 2: The six lines of action of the EOSC model. Source: European Commission 2018a: 9.

The "Architecture" section makes clear that the EOSC, implemented as a federation of infra-structures, should provide *"the solution to the current fragmentation in research data infrastructures which are insufficiently interoperable"*. Regarding "Data" the core role of the FAIR data principles is highlighted. It is also worth noting that the expert group of the Open Science Policy Platform, which advises the European Commission on open science policies, recommended that the EOSC *"must also support the diversity of requirements across all disciplines"* and *"be compatible with other related initiatives including national, European and Global Research infrastructures to ensure interoperability and the free movement of information across all national and international boundaries and between disciplines, while being sensitive to ethical, societal and legal issues"* (OSPP 2018: 8).

The EOSC initiative is well on its way, as on 23 November 2018 the initial EOSC portal was launched, enabling access to first sets of service and data resources[6]. But the realisation of the EOSC vision, particularly wide adoption and use by different stakeholders, will require much more than technical services. In June 2019 the EOSC Executive Board published the strategic implementation plan for the EOSC that gives a comprehensive overview of the planned activities which will contribute to the realisation of the EOSC goals in the period 2019-2020 (EOSC Executive Board 2019). It is primarily intended for use by stakeholders engaged in building the EOSC, and includes the list of related Horizon 2020 projects. Connection of most infrastructures and services to the EOSC is scheduled to be reached in the second quarter of 2020 (EOSC Executive Board 2019: 14).

ARIADNEplus and the EOSC

As elaborated by Niccolucci & Richards (2019), alignment and integration with the EOSC is part of the innovation objective and approach of ARIADNEplus, which is "*based on the provision of innovative and advanced web services in a cloud environment, coherent with the vision, and integrated in the implementation of the EOSC. ARIADNEplus will progressively set up an ecosystem for digital archaeological research which incorporates data and services and enables the use of cloud-based Virtual Research Environments (VRE)*". Further, referring to envisaged impact, *"The overall strategy with regard to improved use of resources, economies of scale and cost-savings is Cloud-based virtualisation and integration in the EOSC"* (Niccolucci & Richards 2019: 9 and 23).

ARIADNEplus follows the EOSC vision of providing a wide range of data resources, services and tools on a Cloud-based research platform. Virtual Research Environments (VRE) on the platform will allow researchers to use online tools for different archaeological research tasks and types of data. Providing such online VREs means researchers can invest less effort to acquire and manage such tools. The approach allows cost-savings for the research community, while also creating opportunities for research groups to jointly address research questions.

While this will be enabled by ARIADNEplus VREs on the fully operational D4Science platform[7], integration of ARIADNEplus, and other domain research e-infrastructures, in the EOSC will be established based on catalogues of services and data resources. The optimal way to exchange catalogue information with the EOSC "marketplace" of service and data resources will be investigated. ARIADNEplus employs the CRMpe catalogue model created by the Parthenos project[8], as a general model for all humanities and heritage catalogues. The CRMpe[9] covers activities, actors, procedures, datasets and software. A mapping can be established between this model and the EOSC catalogue modes to enable interoperability.

EOSC innovation activities

- Develop Cloud-based ARIADNEplus data services to allow improved use of resources, economies of scale and cost-savings.
- Cooperate with other e-infrastructures for digital heritage and humanities research regarding harmonisation of data catalogues.
- Create a mapping between the ARIADNEplus catalogue model and the EOSC model for service and data resources, and register resources in the EOSC.

[6] EOSC portal, https://www.eosc-portal.eu

[7] D4Science (CNR-ISTI, Pisa, Italy), https://www.d4science.org

[8] Parthenos, http://www.parthenos-project.eu

[9] CRMpe stands for CIDOC Conceptual Reference Model (CRM) – Parthenos entities (pe).

3.2 Data resources and services

3.2.1 From access to archaeological data to digital archaeology

The core achievement of the first ARIADNE project was the implementation of a fully functional pipeline to harvest, integrate and make searchable records from repositories and databases of institutions located in different countries within a portal data. The pipeline aggregates the data records into the ARIADNE catalogue, and feeds the search portal with records that are integrated based on general standards, e.g. WGS84 coordinates for locations, and vocabularies such as the Getty Art & Architecture Thesaurus, the PeriodO system for cultural periods, and others. The data portal then provides advanced search options such as multi-lingual, subject-based search, map-based search, and timespan-based search with a visual interface for selecting date ranges.

ARIADNE integrated mainly data records from monument and site inventories, excavation archives, fieldwork reports, and artefact databases. ARIADNEplus now aims to aggregate and integrate datasets from a wide range of archaeological domains of research, and provide research tools and services in addition to data search and access services. The Cloud-based D4Science platform of CNR ISTI[10] will power the ARIADNEplus data pipeline as well as provide several enhanced and new services and tools. The existing datasets in the ARIADNE Catalogue will be updated and the pool of data records extended geographically, temporally and thematically by incorporating additional datasets. The records will be integrated using a Linked Data approach that enables novel ways to search and browse data based on detected relations between them.

Furthermore, ARIADNEplus will offer Virtual Research Environments that, in addition to data discovery and access, will provide services and tools for different types of data and tasks. For instance, these will include geo-spatial/GIS data services, tools to annotate texts and images, and Natural Language Processing of documents to extract specific information they contain. Existing services for visual data objects (e.g. 3D models) will be enhanced and new ones developed. Several pilots will be developed to test and demonstrate innovative uses of the new digital research platform.

3.2.2 Incorporating datasets from many research domains

The first ARIADNE project incorporated mainly monument and site inventories, excavation archives, fieldwork reports, and artefact databases. ARIADNEplus aims to integrate a wider range of data types from different archaeological domains of research, and integrate them as far as possible at item level (described in the next section).

The new thematic domains from which datasets will be incorporated in the ARIADNEplus infra-structure include palaeo-anthropology, bio-archaeology, environmental archaeology, maritime and underwater archaeology, standing structures (e.g. ancient buildings), spatio-temporal/GIS data, remote sensing (e.g. LiDAR and satellite data), inscriptions (epigraphy, rock carvings), burial archaeology, inorganic materials study and dating; *Section 7.1* (Annex A) gives an overview of the domains.

This requires standardised description of records of different types of data by the providers, based on application profiles jointly developed by domain researchers, data managers and vocabulary experts. Such profiles allow mapping of the schemas used in domain databases to the common

[10] D4Science (CNR-ISTI, Pisa, Italy), https://www.d4science.org

CIDOC Conceptual Reference Model (ontology)[11] and appropriate extensions, e.g. CRMarcheo for archaeological excavation or CRMba for standing structures as studied in buildings archaeology.

The application profiles also recommend general and domain-specific vocabularies for metadata elements, for example, for locations WGS84, for cultural periods PeriodO, for subjects the Getty Art & Architecture Thesaurus and more specific objects and scientific vocabularies.

Based on the application profiles the data records can be aggregated and integrated for use with different services and tools. For instance, Linked Data based search, visualisation and exploration of data maps (e.g. finds distribution maps based on location data), and item-level access and study with data-specific tools (e.g. for 3D models).

Data domains innovation activities

- Integrate data from a wider range of archaeological research domains, e.g. palaeoanthropology, bioarchaeology, environmental archaeology, scientific analyses (e.g. material sciences), remote sensing, standing structures, among others.
- Develop application profiles for data records representing different archaeological domains with domain experts, and support data providers in mapping to the ARIADNEplus CIDOC CRM based data model.
- Promote the standardised description of the metadata records with appropriate general and domain-specific vocabularies (e.g. WGS84, PeriodO, Getty AAT, scientific domain terminologies).

3.2.3 Increasing the depth of data integration and using Linked Data

ARIADNE developed a system to aggregate, integrate and search federated data resources held by different data providers. ARIADNEplus will integrate more data resources and, where possible, at the item-level rather than only through collection-level access, as in the case of some ARIADNE recources. What does collection- and item-level access mean, and what will item-level access enable?

At present over 1.9 million data records are integrated in the ARIADNE catalogue and portal. These provide access to about 3.7 million data items, as in many cases one record describes and directs the portal user to data sets of hundreds or thousands of items of fieldwork archives, artefact databases, scientific databases (such as dendrochronology data, etc.) that are accessible in a repository or specialised database. Thus there are data collections from which each item can be found directly within the portal, while in other cases only indirectly by following a link in the record of the collection served by the portal.

This difference between item-level and collection-level access is due to the technical setup of some data collections which make it difficult to provide records of individual items. In other cases it is preferable to provide access at a higher level, e.g. the description of a database or group of items in a repository, rather than individual items without contextual information. Therefore, in ARIADNEplus for each new data collection and updates of some of the already present collections in the ARIADNE catalogue, the best integration approach will be defined through taking account of the content and technical setup of the collection.

For the data aggregation and integration ARIADNEplus applies a process that can cope with the heterogeneity of data structures and the granularity implemented by archaeological organisations. In this two-step process, first what a data provider makes available as a structured collection is

[11] CIDOC Conceptual Reference Model, http://www.cidoc-crm.org

harvested, e.g. a project database or archive, then the collection is disaggregated to the finest granularity possible, which can include records of single items (e.g. fieldwork reports) or of several items (e.g. different images of an artefact).

In the process, the metadata records are also transformed to Linked Data, which are based on the mapping of the collection to the ARIADNEplus data catalogue model (an implementation of the CIDOC CRM ontology) and the vocabularies used to describe metadata elements, e.g. subjects, locations, cultural periods (see previous section). The integration of the Linked Data in the *ARIADNEplus Content Cloud (AC)* then allows using semantic search functionality to discover and retrieve data resources based on semantically defined relations between them. This provides the potential to search and access the *AC* data for specific research questions.

Experts agree that Linked Data comply with most of the FAIR data principles[12], and the project will of course take care to build the *AC* and its dataset of Linked Data following these principles. Linked Data-based ontologies and vocabularies that are already in the Linked Open Data Cloud[13] form hubs for linking data across providers, and ARIADNEplus uses some of them (see *Section 4.4*). Thus the *AC* can also relate to data from other Linked Data providers, and vice versa. Therefore ARIADNEplus will make its Linked Data accessible to external developers to consume and interlink data.

In summary: The ARIADNEplus data aggregation and integration process takes account of the heterogeneity of data structures and the granularity archaeological organisations have implemented. The resulting *ARIADNE Content Cloud (AC)* is a Linked Data dataset that combines data at any level of granularity, systematically linked together and searchable based on domain ontologies and vocabularies. The *AC* dataset will as far as possible correspond to the FAIR principles. The dataset will also refer to relevant data from other reliable Linked Data providers and be linkable by them.

Data integration innovation activities

- Extend the interoperability of data records integrated from federated collections/databases to item-level access.
- Follow FAIR principles in building the ARIADNE Content Cloud (AC) to create a FAIR dataset of Linked Data.
- Apply Linked Data methods to exploit relations between data within the AC as well as with reliable external data resources.

3.2.4 Extending the portfolio of services for digital archeology

ARIADNEplus aims to provide a range of services for digital archaeology on the Cloud-based D4Science platform. The goal is to enable innovative and effective ways to carry out data-based research in archaeology (e-archaeology). Therefore the portfolio of Cloud-based services offered to users will be extended from data discovery and access to services supporting digital research, analysis and interpretation.

The provision of Cloud-based services ("Software as a Service") can help reduce costs of technology management as it avoids scholars investing effort to acquire, implement, maintain and upgrade software for research. This allows cost-savings for the research community while at the same time opportunities for research groups to jointly use services for addressing research questions.

[12] For example see GO-FAIR: FAIRification Process, https://www.go-fair.org/fair-principles/fairification-process/; Cox & Yu (2018); Hasnain & Rebholz-Schuhmann (2018).

[13] Linked Open Data Cloud, https://lod-cloud.net

ARIADNEplus services will enable much more than data discovery and access. For example, when researchers discover data items of interest they will also be able to study them with available services (e.g. tools for exploring 3D models). ARIADNE already offered some digital research services like the visual media and landscape services, which enable effective online publication and exploration of images (e.g. Reflectance Transformation Imaging - RTI) and 3D models of objects and landscapes. In ARIADNEplus these services will be enhanced and new ones developed, e.g. visualisation in 3D of the layers of an excavation and the related documentation. Also planned is provision of geo-spatial/GIS data services, Natural Language Processing of documents to find and extract specific information, and tools to annotate texts and images, e.g. fieldwork reports, artefact or laboratory images; *Section 7.2* (Annex B) gives an overview of the services.

The interest of archaeological researchers and data managers in using different existing and planned services has already been investigated in an online survey (ARIADNEplus 2019). Among the planned new services respondents appreciated most were services for searching and visualising geo-spatial/GIS datasets. Respondents were also particularly interested in using Linked Data to interlink their own and other datasets. In ARIADNEplus Linked Data standards and technologies will be employed for data integration and some of the search services. Furthermore respondents were interested in using services for working with visual content (e.g. 3D models, LiDAR imagery). Such services were already offered by ARIADNE but not integrated on a service platform, for which ARIADNEplus now employs the Cloud-based D4Science platform.

In the development of the ARIADNEplus services IT specialists and archaeologists will work closely together to ensure that the services support user needs regarding new research capabilities. This will be an iterative process in which archaeologists trial service prototypes and suggest improvements and additions to be implemented, so that archaeologists can generate new research insights with them.

Service innovation activities

- Offer easy-to-use Cloud-based services to make it easier for researchers to acquire and manage software for research ("Software as a Service").
- Bring IT specialists and archaeologists together to jointly develop services that support user needs regarding new research capabilities.
- Provide different data search, visualisation and access services as required for integrated data, e.g. Linked Data-based search, sites and finds distribution maps, item-level access and study with data-specific tools (e.g. for 3D models).

3.2.5 Providing virtual research environments

ARIADNEplus will provide Virtual Research Environments (VREs) for digital archaeology on the Cloud-based D4Science platform. In addition to data discovery and access such environments combine more specific services and tools which research communities can use for different tasks and types of data.

In ARIADNE the state of e-archaeology in different fields of research, perceived difficulties, and requirements for progress towards innovative solutions have been investigated (ARIADNE 2017b). The study results suggested there is much potential for ARIADNE to provide VREs, with the proviso that the data infrastructures and services will have to take account of the multi-disciplinarity of archaeological research, particularly different data standards and vocabularies being used by different research communities.

Advanced VREs make tools available in a highly integrated way to support research workflows. Finding out which tools to combine in VREs will be explored with researchers in workshops and experimental implementations on the D4Science platform. VREs will also be used in pilots undertaken by project partners to demonstrate the innovative capabilities of ARIADNE services and data for archaeological research as well as public/community archaeology.

VRE innovation activities

- Set up an ecosystem for digital archaeological research that incorporates data and services and enables use of Virtual Research Environments (VREs).
- Provide Cloud-based VREs that combine services as needed by researchers of different thematic areas.
- Demonstrate innovative uses of the new research platform and data resources in a number of pilots.

3.3 Taking account of user needs

The innovation strategy takes account of the results of the ARIADNEplus community needs survey (ARIADNEplus 2019). A broad survey on the needs of the ARIADNEplus user community regarding data sharing, access and (re)use, new services, and related training needs was carried out.

The survey results are based on 484 questionnaires received from all 27 ARIADNEplus partner countries and a few other countries. 93% of the respondents are professionally based in a European country; 46% said the organisation they work for is a member of the ARIADNEplus consortium.

The organisational background of most respondents is a university or public research organisation (53%), museum (19%), governmental institution (15%) or a private company or research institute (8%). Regarding professional activities, 53% are archaeological researchers (field work), 9% laboratory-based researcher, 13% managers of an institutional repository or other data access services, 7% managers of project databases, 7% directors of an archaeological institute or research centre/laboratory, 12% other (various academic, technical and data management activities).

Results of the survey were compared, where possible, to those of the ARIADNE survey in 2013 (ARIADNE 2014). The analysis of the results focused on the match between the perceived user needs and the planned ARIADNEplus technical and other services, and suggestions were provided for activities likely to enable an optimal match.

3.3.1 Data publication

In the ARIADNE/plus surveys we used the concept of “data publication” to emphasise the common understanding that publication means that the data indeed is publicly available. Researchers often share data directly with colleagues but do not make them publicly available (e.g. in a repository). This means valuable data remains within small circles of peers and is not available to other researchers and the wider public. Moving more data from closed-circle or not shared to “open data” requires overcoming strong barriers, which are addressed in the next section.

Survey results

In the 2019 survey, around 65% of respondents said they published some to all of their project data through an accessible repository, in 2013 around 50% said they did so. The results suggest that from 2013 to 2019 in the ARIADNE/plus communities of respondents the sharing of data through accessible repositories increased significantly by up to 15%. The percentage of 65% in 2019 is over

30% more than reported by other surveys with respondents from different disciplines, e.g. Tenopir et al. (2015) 30%, Figshare (2018) 33%.

Many survey respondents share data as supplementary material to publications and reports, 81% of 449 respondents in 2019 (82% of 520 respondents in 2013). In 2019, 13% said they do so in all or most, 25% in many, and 42% at least in a few projects. In other surveys the percentages were smaller, Tenopir et al. (2015) 19.4%, Figshare (2017) 34%.

The explanation for the much higher percentages in the ARIADNE/plus surveys could be that many of the respondents are obliged to provide fieldwork reports to a national heritage authority, and do this with supplementary material added.

The survey participants were also asked if they agreed with the statement: *"In the last 5 years the readiness of archaeologists to share data through publicly accessible repositories or databases increased"?* – 83.2% of 376 respondents agreed. However, several respondents perceived a higher awareness among archaeologists that data should be made available, but little increase in readiness to do so. Others felt that the increase is taking place only slowly. More has to be done to foster data sharing.

Suggestions for ARIADNEplus

While the ARIADNEplus 2019 survey shows good results for sharing reports and data through institutional repositories (e.g. repositories of heritage authorities or research centres), many archaeologists in European and other countries do not have available yet a state of the art digital repository for archiving and sharing their data.

This issue is being addressed by the COST Action Saving European Archaeology from the Digital Dark Age (SEADDA). SEADDA and ARIADNEplus share the goal of making archaeological data FAIR (Findable, Accessible, Interoperable and Reusable), especially by supporting knowledge exchange and collaboration within data repositories and e-infrastructure.

The core requirement for moving research data into accessible repositories is decisive open data mandates by research funders, coupled with funding of the basic costs of domain repositories and researcher data deposition costs (e.g. as part of research grants).

Suggestions to increase further the sharing of archaeological data through appropriate repositories are:

- Continue the good collaboration between ARIADNEplus and SEADDA on capacity building for new repositories and use of the ARIADNEplus digital infrastructure
- Support strict open data policies of funding bodies and institutions – providers of data repositories and e-infrastructures should give full support to such mandates.

3.3.2 Barriers to data deposition & sharing

Core functions of the ARIADNEplus digital infrastructure are to aggregate data from archaeological repositories and provide search and access services. Therefore the ARIADNEplus initiative depends on repositories richly filled with accessible data shared by researchers. It cannot ignore obstacles which hinder researchers in sharing their data in an open manner. Rather the initiative must support researchers in data sharing and help ensure that they receive appropriate credit for doing so.

Survey results

In the ARIADNE/plus surveys 2013 and 2019, participants were given a list of potential barriers for researchers to deposit their data in digital repositories and share it with others. The barriers which

respondents perceived as most critical were the same, albeit with some differences regarding the percentages of "very" or "rather" important combined:

- *A lack of professional recognition and reward:* was considered as most critical by 75.5% of respondents in 2019, while 72% in 2013.
- *The work effort for providing the data and metadata in the required formats:* was an important barrier for 74% of respondents in 2019, while in 2013 more respondents worried about the work effort for metadata and data (80%).
- *Intellectual property rights issues:* was a concern for 75% of respondents in 2019, while significantly less in 2013 with 65%.
- Two barriers were perceived as somewhat less important with about the same percentages: *Lack of appropriate repositories* with 67% in 2019, while 66% in 2013; *the cost for depositing data in a repository* with 59% in both years.

Regarding professional recognition and reward for data sharing, in the latest Figshare survey (2018) the majority of respondents felt that they did not get sufficient credit for data sharing, 58%, compared to 9% who felt they do; 33% were not sure. Fecher et al. (2015) in a large survey with respondents from different disciplines found that *"if I were cited in publications using my data"* would motivate 79.3% of 1,420 respondents to make data available (9.5% said it would not, and 11.2% were undecided)

Suggestions for ARIADNEplus

Understanding obstacles to data sharing and helping to remove them is essential for research data infrastructures such as ARIADNEplus, as well as the underlying digital repositories. Advocates of open data argue that such data will often be (re)used and cited, bringing recognition and rewards to data publishers (incl. data repositories). The scenario is that data citations indicate and acknowledge providers of valuable data, promote further data sharing and (re)use, and enable the impact of open data to be tracked and measured. Most importantly, it would drive the emergence of an academic credit system that appropriately rewards open data sharing.

- Research infrastructure components, protocols and metrics for data citations are in development. ARIADNEplus should investigate how services of the research infrastructure could help identify and track (re)use of data based on data citations (e.g. article-data links) and other indicators.
- As a general requirement for identifying data (re)use, the project could promote and support standardisation of data citation in the archaeological sector, i.e. how data should be cited in publications to ease the identification and tracking of data (re)use.

3.3.3 Reuse of data

Sharing data is important but without (re)use the benefits associated with open data sharing would not materialise. There are many good arguments for making data available, for instance, that reported research results can be scrutinised and duplicative data collection prevented. Particularly strong however is the argument that reuse of data, for example to investigate new research questions, allows exploitation of previous investment. Preserved data that is being reused gains in value, otherwise it might be perceived only as a cost factor.

"Return on investment" expected by research funders explains much of the increasing pressure on researchers to share their data from publicly funded research for reuse. It is also very important for repositories to document not only downloads but actual reuse. Metadata with rich context

information is essential for reusing data, as is a license that clearly states what users are allowed to do with the data.

Survey results

Results of the ARIADNEplus 2019 survey confirm that archaeological researchers often (re)use available data and allow some insights into what and how that data will be (re)used. The survey participants were asked, *"Did you / your research group in the last 2 years use any data which other researchers made available through a publicly accessible digital repository or databases?".* – An astonishing number of 220 respondents said they did and also briefly described the data types and/or the sources.

In comments 34 respondents also gave reasons why they did not (re)use other's data. Most said relevant data was not available or posed some problems, difficulty to access or use, missing licensing information, lack of support, among others. Some also said that they did not need data from other researchers.

The 220 respondents were asked "*What was the main purpose of the data reuse?"*, and three predefined purposes and the option "Other" offered. Building a database for the research community was a purpose for 31%, comparison to own results for 55%, and use together with own research data for 63% (multiple answers were possible). Few mentioned other purposes for the data (re)use, for example, to use it as test data (e.g. "test algorithms and approaches") or as a conceptual resource (e.g. "typology terms", "data structure").

Suggestions for ARIADNEplus

- ARIADNEplus should promote and support data reuse so that the investment in the collection of archaeological data can be exploited further for research, education and other purposes. Ways to enable easy reuse of data with the ARIADNEplus infrastructure should be investigated.
- Different purposes and forms of data reuse should be considered to better understand actual practices of data reuse in archaeology so that these can be supported effectively.

3.3.4 Data search & access services

ARIADNEplus will incorporate data from a wider range of archaeological research domains than ARIADNE, including environmental archaeology, maritime and underwater archaeology, biological and inorganic materials studies, radiocarbon, dendrochronology and other dating methodologies, among others. Furthermore, the project aims to integrate more datasets at item level to provide advanced semantic data search to find data items based on semantically defined relations.

The ARIADNEplus 2019 survey investigated three closely related questions on data search and access: including current online availability of the different types of archaeological data, and how helpful it would be to discover and access the data via the ARIADNEplus portal at both the collection level and item level.

Survey results

Online accessibility of data types:

Survey respondents were asked to rate the current availability of the different types of data ARIADNEplus aims to mobilise and integrate into the dataset catalogue and portal. The data types were rated as follows:

- good availability: archaeological sites and monuments data (usually provided by heritage authorities), national GIS data and maps (from mapping agencies), and satellite or airborne remote sensing data (in Europe offered freely by the European Space Agency);
- less good availability: data and documentation of fieldwork (excavation, field survey/prospection, fieldwork reports), and databases and catalogs of various artefacts (e.g. museum collections);
- poor availability: dating data (e.g. dendrochronology, radiocarbon) and scientific data/analysis of biological and inorganic remains. Also the availability of environmental archaeology and maritime & underwater archaeology data was perceived as poor.

ARIADNEplus portal for data discovery & access

One particularly important objective of the survey was to identify if respondents perceive support by the ARIADNEplus portal to discover and access more helpful for some data types than for others. A comparison of the online accessibility rating of the data types and the helpfulness of portal support for discovering and accessing datasets or collections of such data showed:

- surprisingly, the appreciation of support was lower for data types for which the accessibility was evaluated as insufficient;
- respondents were most appreciative of portal support for discovery and access of sites and monuments databases or inventories, national GIS data & maps, and satellite or airborne remote sensing data (e.g. LiDAR), although the online accessibility of these data types was rated much better than that of other data types;
- the rationale for ARIADNEplus should not be to prioritise support for data types which are already much more accessible than others; the fact that the more accessible types are being provided by national mapping and heritage authorities indicated ARIADNEplus should prioritise other data types.

Suggestions for ARIADNEplus

The survey results tentatively suggest the following priorisation of data types for mobilisation and integration in the ARIADNEplus portal:

- Data types with high or medium appreciation of portal support, and currently medium or low online accessibility. These types are ranked according to the appreciation of support and level of accessibility:
    - Excavation data (e.g. excavation archive)
    - Artefact/finds databases or image collections
    - Radiocarbon, dendrochronology and other dating data
    - Environmental archaeology datasets
    - Unpublished fieldwork reports
    - Field survey/prospection data
- Next would come the following data types with lower appreciation of respondents of portal support:
    - Maritime and underwater archaeology data
    - Scientific data/analysis of inorganic remains
    - Scientific data/analysis of biological remains

- Inscriptions, coins or other special databases

But it must be noted that the lower appreciation is very likely due to fewer survey participants being specifically interested in these data types (e.g. martime & underwater archaeologists).

The survey participants also rated which data types they would find helpful for their research if able to search items within datasets integrated from multiple sources (item-level access). The results confirmed the evaluation above, except that artefact/finds databases or image collections were ranked highest.

3.3.5 Special services for researchers and data managers

A wide range of enhanced and new services for researchers and data managers is foreseen to be provided on the D4Science platform for virtual research environments (an overview is given in *Section 7.2*, Annex B). Therefore an important goal of the ARIADNEplus 2019 survey was to find out which ones the respondents perceive as particularly helpful for their research or data management. The services in question are for end-users, "back-office" services, those which end-users do not use directly were not included in the survey.

Survey results

A very encouraging survey result is that respondents appreciated very much services which are already available (and which they may have already used): Register a dataset in a portal that allows the searching of data from many providers; Discover & access archaeological data stored in repositories in different European & other countries; Spatially and/or chronologically defined search options.

Services for searching and visualising geo-spatial/GIS datasets were the highest ranked among the new services, and are part of the plan of services ARIADNEplus will implement. Respondents were also particularly interested in using Linked Data to interlink their own and other datasets. Applying Linked Data standards and technologies is the general approach in ARIADNEplus for data integration and some of the search services. Project datasets in Linked Data formats will also be made available (e.g. via an API) to external developers for interlinking datasets.

Furthermore, respondents were interested to use services for working with visual content (e.g. 3D models, LiDAR imagery), considerably more than using services for textual content. Visual content services were already offered by ARIADNE but not integrated on a service platform, for which ARIADNEplus will employ the D4Science platform.

Lowest on the list is a service for mapping databases to the CIDOC-CRM extended for archaeological research data. This result did not come as a surprise because the service is specifically for data managers (databases, repositories) and these made up only 20% of the survey respondents.

Suggestions for ARIADNEplus

The main suggestions that can be derived from the survey results are:

- Devote special attention to the new services for search and visualisation of geo-spatial/GIS datasets.
- Prioritise the use of Linked Data for interlinking datasets, particularly at item level.
- Continue to enhance existing, and develop new visual content services of interest.
- Evaluate further which services for textual content are of interest to users, including services not yet considered.

- Promote use of the CIDOC-CRM by making clear its capability to integrate research data conceptually, especially regarding the ontology extensions developed in the ARIADNE project for archaeology (e.g. excavations, standing structures, epigraphy).

3.3.6 Training needs

ARIADNEplus training for researchers and data managers of archaeological projects will be provided in the Transnational Access (TNA) programme as well as at domain conferences in the form of tutorials and short courses, e.g. on how to use ARIADNEplus services for research and data management. In matters pertaining to archaeological repositories ARIADNEplus will coordinate activities with the SEADDA project. SEADDA aims to foster the development of archaeological data repositories in countries where the research community lacks an appropriate repository.

Survey results

Training in the application of the FAIR data principles in archaeology would be appreciated most, both by researchers and data managers, and ARIADNEplus is committed to support these principles within the archaeological sector.

Significantly less appreciation was expressed for training in how to create and implement a data management plan (DMP), manage a digital repository, produce metadata and use domain vocabularies to describe archaeological datasets. Data science skills, managing datasets of a large archaeological project, and depositing project datasets in a digital repository were scored higher.

That researchers were the largest group in the survey sample had a considerable impact on the results. Researchers worried about additional data-related work, which explains why training regarding DMPs, metadata and vocabularies is appreciated less.

Inconsistently, training on data deposition appeared to be welcome, despite the (not recognised) fact that this would require dealing with metadata and vocabularies. Awareness of an increasing expectation that data from funded research projects should be deposited may have contributed to this result.

Suggestions for ARIADNEplus

- *Application of FAIR data principles:* Support FAIR data policies and good practices of researchers and data managers through guidelines, workshops and training (e.g. online webinars to reach many researchers). Training offers should be as practical as possible, distinct from the broad wave of general information on the FAIR principles; focus on what matters for archaeological researchers and data managers specifically.
- *Data science skills:* ARIADNEplus has limited capacity to raise the level of data science skills of archaeological researchers, i.e. use of advanced data processing and analysis methods. This should be done by dedicated courses within universities and data science centres. However, a focus on data science skills is possible, related to the ARIADNEplus Cloud-based Virtual Research Environments (VREs).
- *Research Data Management (RDM):* Continue making researchers aware of available guides to good practice, e.g. the guides offered online by Archaeology Data Service and Digital Antiquity for different types of research data.
- *Data Management Plans (DMP):* To define and implement a DMP and related activities (metadata, vocabularies) adds work, and researchers are unsure they will benefit from this additional work. Case studies making clear the benefits could promote more interest in data management planning for archaeological projects.

4 The WP6 task forces

The innovation strategy covers activities and contributions of all project work packages. But for some topics, which require special attention, small groups of partners in WP6 are acting as "task forces" to provide knowledge and guidance on how to promote innovation and impact.

4.1 Overview of topics

The four WP6 task forces are addressing the following topics defined in the project work plan under Task 6.2 "Fostering innovation by ARIADNEplus, stakeholders and users":

- e-Infrastructure synergies and joint development
- Internationalisation of the ARIADNEplus approach
- Benefits of Linked Data technologies
- Innovating services with industries

These are very different topics and, accordingly, the focus and activities of the task forces are different. In the sections that follow for each the focus, background, and activities and results so far are described, and a summary and outlook is given.

4.2 e-Infrastructure synergies and joint development

Focus of the task force

The focus of this task force is described as follows (project Task 6.2.1 description): *"Targeted activities are required to fully exploit synergies and complementary capabilities among research e-Infrastructures, thereby raising their innovative impact and economies of scale. Building on achievements of the digital humanities cluster project PARTHENOS (coordinated by PIN), ARIADNE-plus will promote this agenda further together with new entries such as the E-RIHS heritage science infrastructure. The overall innovation goal here is to integrate humanities and heritage e-infra-structures as part of the European Open Science Cloud (EOSC)."* – Task lead: PIN, six other assigned partners.

Background

Research e-infrastructures and services are important pillars and drivers of collaborative and data-intensive research. They provide researchers access to distributed, but shared digital resources (data, services, tools), and can support advanced and innovative research across institutional and national as well as disciplinary boundaries.

In Europe, considerable investment has already been made to implement generic and discipline-specific research e-infrastructures, but the e-Infrastructure Reflection Group (e-IRG), the main advisory body for European research e-infrastructures, regulary noted fragmentation, lack of coordination and interoperability between the existing e-infrastructures. The e-IRG introduced the concept of e-Infrastructure Commons, and asked all stakeholders to increase coherence, inter-operability and resource sharing. These are required to enable synergies, cost-effectiveness and sustainability in supporting ICT-enhanced research across disciplines (e-IRG 2012; e-IRG 2013; e-IRG 2016). So far, however, interoperability and sharing of resources between different research e-

infrastructures has not been sufficiently realised (e-IRG 2019).[14] The European Open Science Cloud (EOSC) initiative now aims to provide a basis for tighter integration of research e-infrastructures, services and data.

The integration of ARIADNEplus, and other domain research e-infrastructures, with the EOSC will be realised through catalogues of services and data resources. ARIADNE was a front-runner in establishing a standards-based catalogue of research datasets. The ARIADNE Data Catalogue Model has inspired the model of the humanities e-infrastructure cluster project PARTHENOS, which developed a model with additional entities (Frosini et al. 2018). These are now also included in the ARIADNEplus model.

Therefore harmonisation of catalogues of heritage and humanities research resources (data, services) at the European and national levels is possible in principle, and would fulfil the FAIR criteria of findable and accessible for the catalogued resources. In addition, more complex measures are necessary to make the resources interoperable and reuseable, through the EOSC or directly between research infrastructures. This requires working towards common standards regarding data, technologies and policies to enable the expected synergies between research e-infrastructues and, more importantly, value-added for the research communities.

Activities and results

Since the start of ARIADNEplus, partners of the task force promoted coordination and collaboration with initiatives and projects in which they participate, as well as others. The following examples illustrate the different levels, types and results of the activities.

The Group of European Data Experts (GEDE)[15] is an interest group of the Research Data Alliance with a focus on FAIR digital objects. GEDE involves experts from many research e-infrastructures at different levels. GEDE has been asked officially to contribute to the EOSC development process. The ARIADNEplus coordinator F. Niccolucci (PIN) has participated in GEDE since 2016 and is now a member of the GEDE EOSC core group.

The RISCAPE project[16] was in charge of mapping the international landscape of research infra-structures. PIN has provided the perspective of digital cultural heritage and archaeology to their final report (RISCAPE 2019).

The expertise brought together in ARIADNEplus is recognised widly in the cultural heritage sector. For example, at the Europeana Conference 2019 a workshop on what the EOSC offers to the sector as well as a session on how to implement the FAIR principles were both organised by ARIADNEplus partner MIBAC-ICCU and supported by the project coordinator.

The European Research Infrastructure for Heritage Science (E-RIHS)[17], which has been on the ESFRI Roadmap since 2016, brings together research centres with a focus on material analyses of heritage objects. ARIADNEplus partners contribute to their data stratey, i.e. application of quality concepts such as the FAIR principles and repository certification for data resulting from scientific conservation analyses.

[14] Some general distinctions are between e-infrastructures offering their services towards all communities/ disciplines (called “generic” or “horizontal”) and community/discipline-specific e-infrastructures (called “thematic” or “vertical”); both can be positioned at one or more levels, local/regional, national, European, international/global.

[15] GEDE - Group of European Data Experts in RDA, https://www.rd-alliance.org/groups/gede-group-european-data-experts-rda

[16] RISCAPE - European Research Infrastructures in the International Landscape, https://riscape.eu

[17] E-RIHS - European Research Infrastructure for Heritage Science, http://www.e-rihs.eu

The Distributed System of Scientific Collections (DiSSCo)[18], which has been on the ESFRI Roadmap since 2018, is comprised of research centres in the field natural history, geo- and bio-diversity. PIN contributed to a DiSSCo networking workshop on research data management.

Social Sciences and Humanities Open Cloud (SSHOC)[19] is a H2020 project (2019-2022) that involves over 40 organisations. SSHOC aims to interconnect existing and new infrastructures from the social sciences and humanities and foster interdisciplinary collaboration and research. ARIADNEplus partners also participate in SSHOC, for example, with contributions on archaeology and heritage FAIR and Linked Data (ADS, CNR, DANS).

EOSC-Pillar[20] coordinates national Open Science efforts across Austria, Belgium, France, Germany and Italy, and supports their readiness for contributions to the EOSC. An example of technological knowledge transfer here is the reuse of the Natural Language Processing tool developed in ARIADNE. In EOSC-Pillar, the NLP tool will be applied to the digital results of diagnostic material analyses. ARIADNEplus supports this by providing the necessary information to adapt the tool to a Cloud environment.

The Italian Computing and Data Infrastructure (ICDI)[21] initiative, led by the GARR High Speed Research and Education Network, brings together all Italian institutions involved in ESFRI infrastructures and e-Infrastructures. ARIADNEplus is a member of the initiative and actively participates in the definition of the national roadmap for EOSC participation.

Summary and outlook

According to its overall goal the task force supports the integration of e-infrastructures in the fields of humanities, heritage and archaeology in the EOSC. The activities concern contributions to integration within the field as well as in the EOSC.

The focus so far has been on European level expert consultation, projects and initiatives. However, the mentioned Italian Computing and Data Infrastructure (ICDI) initiative is not a singular case. For example, CNRS (France) is present in ARIADNEplus with the national HumaNum[22] digital humanities infrastructure, which includes an archaeological consortium, MASA - Mémoire des Archéologues et des Sites Archéologiques.[23] Partners from other ARIADNEplus countries (ADS, CNR, USW) are scheduled to contribute to a forthcoming MASA conference on FAIR Heritage (May 2020).

Members of the task force play a key role also in other countries, for example DANS (Netherlands) as the core social sciences and humanities archiving and networked service provider, ATHENA RC (Greece) as a leading e-infrastructure centre at the national level as well as in European projects (e.g. OpenAIRE).

[18] DiSSCo - Distributed System of Scientific Collections, https://www.dissco.eu

[19] SSHOC - Social Sciences and Humanities Open Cloud, https://sshopencloud.eu

[20] EOSC-Pillar, https://www.eosc-pillar.eu

[21] ICDI - Italian Computing and Data Infrastructure, https://www.icdi.it/en/

[22] HumaNum - Très Grande Infrastructure de Recherche des Humanités Numérique, http://www.huma-num.fr

[23] MASA - Mémoire des Archéologues et des Sites Archéologiques, http://masa.hypotheses.org

4.3 Internationalisation of the ARIADNEplus approach

Focus of the task force

The focus of this task force is described as follows (project Task 6.2.2 description): *"Open Innovation, Open Science, Open to the World" (Carlos Moedas, Commissioner for Research, Science and Innovation) will be guiding principles of ARIADNEplus. The e-Infrastructure capabilities are being extended with an international perspective, already involving partners from Israel, Japan, Argentina and USA. This will allow transfer of knowledge and technological solutions for adoption also by other organisations of the world regions addressed. The task will foster in-depth understanding and taking account of different contexts and requirements for adoption."* – Task lead: AIAC, five other assigned partners.

For this task force a special focus on the Middle East and North Africa (MENA) region has been chosen. Reasons for this focus are the rich heritage of Europe's neighbours around the Mediterranean, and institutions in the region could benefit through collaboration with ARIADNEplus through knowledge exchange and around technical services for archaeological data. Other world regions, especially where ARIADNEplus has collaboration in a country (Argentina, Japan, USA), are included in another work package (WP2).

The task leader Associazione Internazionale di Archeologia Classica - AIAC (Italy) has contacts with institutions in the MENA region and participates in projects that have a regional focus. Partners located in non-European countries can serve as examples that the project is extending beyond Europe and takes account of different regional contexts and requirements for adopting the ARIADNEplus approach. However, parts of the MENA region present collaborative difficulties.

Background

The past ten years have been a disruptive and difficult decade for archaeology in the MENA region. The "Arab Spring" in North Africa and countries of the Middle East, political turmoil, terroristic threats and warfare have either terminated or seriously disrupted archaeological fieldwork in several countries in the region (see e.g. Casana 2013 [on Libya and all affected areas]; the special issue of Near Eastern Archaeology, 78.3/2015; Abdulrahman 2017 [Syria]). At the same time, these countries have become vulnerable to the destruction and looting of their cultural heritage. The priority of the international efforts were, and still are, to work with, where possible, local cultural heritage professionals to prevent or at least mitigate destruction and looting (Munawar 2019).

In addition, other support such as monitoring of endangered sites with remote sensing (e.g. Parcak 2017; the EAMENA initiative described below), databases of antiquities to help curtail the sale and acquisition of stolen objects (e.g Libyan Antiquities at Risk[24]), and opportunities for archaeologists and heritage practitioners to gain relevant skills abroad; training in archaeological science and conservation, for instance. It has also been suggested to focus more on analysing excavated but not studied material stored at institutions of less affected countries and in Europe, digitise and bring online archival collections of archaeological investigations conducted in the region for decades (e.g. di Lernia 2015; Mitchell 2019).

The digitisation of cultural heritage in the MENA region presents a mixed picture. There are some outstanding centres with significant and well-maintained digital resources, for example, in Israel the National Library[25], also coordinator of the A-Z Archives Network Israel[26]; in Jordan the MEGA Jordan

[24] Libyan Antiquities at Risk, https://laar.le.ac.uk

[25] National Library of Israel, collections, https://web.nli.org.il/sites/nli/English/collections/

[26] A-Z - The Archives Network Israel Project, http://www.a-z.digital/en/about/

monuments database[27]; in Egypt the Bibliotheca Alexandrina and the Center for Documentation of Cultural and Natural Heritage (CULTNAT)[28]; and some examples of digital libraries like the Qatar Digital Library[29] or the Bibliothèque Numérique Marocaine[30].

Where collections have been digitised in the region the focus often was on Islamic manuscripts held by libraries and archives (e.g. Ghamouh & Boulahlib 2015 [Algeria]; Ghali 2015 [Egypt]; Zaccaria 2018 [Horn of Africa region]; see also Riedel 2016). However the digitised collections often are not accessible online or become inaccessible due to lack of appropriate IT systems, digital curators, issues of ownership or political interference.

The situation is the same in many other African countries where the heritage institutions often even lack a digital catalogue of their physical collections so that scholars could at least know what they hold and may be studied in a reading room at the institution. The whole digitisation process up to web-accessible collections is still too costly, and staff skilled in such work are not available (Asogwa 2011). Digitisation projects usually need to train local staff on the digitisation tasks first, of which the production of metadata is the most demanding (Kilmurry 2019). Where staff have been trained, often they cannot be retained due to low salaries and status of the profession (Garaba 2015). In general, open access repositories for academic research are not well developed (Agabirwe 2018; Piron et al. 2017), and have been hampered by limited state funding and capacity development.

Activities and results

A collaboration has been initiated with the EAMENA project of the Universities of Oxford (lead), Durham and Leicester that is being supported by the Arcadia Fund (UK charity) and the Cultural Protection Fund (managed by the British Council).[31] Since 2015 the EAMENA initiative maps and documents at-risk archaeological sites in the region using airborne and satellite imagery and fieldwork reports (Rayne et al. 2017). Since 2017 the initiative also trains regional heritage professionals in such documentation und population of the EAMENA database.

The EAMENA database is intended as *"a resource that can be used online by researchers and heritage professionals across the MENA region, Europe and beyond"*.[32] It is based on the open source Arches[33] platform designed by the Getty Conservation Institute and the World Monuments Fund. Arches supports standards such as Dublin Core, CDS[34], Open Geospatial Consortium (OGC) standards, and uses a data model based on the CIDOC Conceptual Reference Model (Zerbini 2018). This allows linking based on CIDOC CRM and vocabularies in Linked Data formats. The director of the EAMENA project has expressed their interest in a detailed discussion on Linked Data which, however, will depend on a secured further development of the initiative.

Another ongoing initiative is the North African Heritage Archive Network (NAHAN)[35] which aims to digitise and bring online archival collections of archaeological investigations conducted in North

[27] MEGA Jordan, http://www.megajordan.org

[28] Bibliotheca Alexandrina, https://www.bibalex.org; CULTNAT - Center for Documentation of Cultural and Natural Heritage, http://www.cultnat.org

[29] Qatar Digital Library, https://www.qdl.qa/en

[30] Bibliothèque Numérique Marocaine, http://www.bnrm.ma/bnrm/fr/bibliotheque/nos-collections.html; the Ministry of Culture also avails of an online inventory of material cultural heritage, including archaeological sites, http://www.idpc.ma

[31] EAMENA - Endangered Archaeology of the Middle East and North Africa, https://eamena.arch.ox.ac.uk

[32] EAMENA database, http://eamena.arch.ox.ac.uk/resources/database-2/

[33] Arches, https://www.archesproject.org

[34] International Core Data Standard for Archaeological and Architectural Heritage (CDS)

[35] NAHAN, http://www.nahanweb.org

African countries for decades. NAHAN has been initiated in 2016 by AIAC, supported by ICCROM[36], and involves archives in Europe and North African countries (e.g. Algeria, Marocco, Tunesia). The online catalogue and other services for the digitised collections will be provided by the ChronoCarto platform.[37]

The digitisation of the content of the archival finding aids, inventory cards or forms, and of the actual archaeological documentation (e.g. manuscripts, fieldwork diaries, drawings, photographs) will take many years to come. Regarding archives in Europe we may expect some online accessible collections in 3-5 years, for archives in North Africa the horizon is very likely longer.

Summary and outlook

The focus of the task force is on the Middle East and North Africa (MENA) region in view of providing ARIADNEplus knowledge and technical services for data from archaeological sites and archives. The liaison with EAMENA has good prospects regarding integration of data, based on a Linked Data approach. As in other cases of such initiatives this will depend on secured funds for the continuation of EAMENA during the current period of ARIADNEplus. A workshop led by AIAC is planned in Oxford to work on the interoperability of the EAMENA database.

The focus of the NAHAN initiative is currently on countries in the western part of North Africa. Here instititutions expressed interest to participate, e.g. in Marocco the national Institut des Sciences de l'Archéologie et du Patrimoine (INSAP) and the Centre Jacques Berque, which is a mixed CNRS and Ministry of Foreign Affairs institute. Key archives in France are interested to digitise relevant collections, for example the Archives Poinsott of the Institut national d'histoire de l'art (INHA) in Paris[38], and the Archives nationales d'outre-mer[39] at the Centre Camille Jullian, a joint unit of the Université d'Aix-Marseille and CNRS.

The digitisation of archival inventory and actual archaeological documentation can take several years. In the case of archives in Europe some online accessible collections may be expected in 3-5 years, for archives in North Africa it could take much longer. One new initiative includes the support for an MA degree in information management at the Université d'Aix-Marseille for a Moroccan archaeologist, who will then return to take over the management of the INSAP archives.

4.4 Benefits of Linked Data technologies

Focus of the task force

The focus of this task force is described as follows (project Task 6.2.3 description): *"Linked Data technologies allow standardisation, enhanced interoperability and re-use of open data resources and vocabularies. Uptake by sector organisations and projects so far has been low, with ARIADNEplus being a lead user at the level of sector research infrastructure. Therefore the benefits must be evidenced and communicated thoroughly. For instance, this will include calculating the benefits of applying extensions of the CIDOC CRM with regard to domain and cross-domain data integration, and of the faster integration enabled by the services, expertise and training provided by ARIADNEplus. The results and guidance on how to benefit from Open Linked Data and technologies will be presented at sector events as well as disseminated through other dissemination channels."* – Task lead: USW, three other assigned partners.

[36] ICCROM - International Center for the Study of the Preservation and Restoration of Cultural Property, https://www.iccrom.org

[37] ChronoCarto (AOROC Archéologie et philologie d'Orient et d'Occident, Paris), https://www.chronocarto.eu

[38] Archives Poinsott, https://agorha.inha.fr/inhaprod/ark:/54721/0056534

[39] Archives nationales d'outre-mer, http://www.archivesnationales.culture.gouv.fr/anom/fr/

Background

Linked Data enable interoperability of dispersed and heterogeneous information resources, allowing the resources to become more discoverable, accessible and re-useable. In the fragmented data landscape of archaeology this is substantial value proposition. Use of Linked Data standards and technologies is a core approach of data integration in ARIADNEplus as it enables novel ways to search and browse the data based on detected relations between them. The approach also allows linking ARIADNEplus data with relevant other datasets in the so-called Linked Open Data Cloud.[40]

At present there are only few such resources in the LOD Cloud and linked together through the use of common vocabularies. The good message is that some vocabularies which are relevant for ARIADNEplus are already there, e.g. Art & Architecture Thesaurus, PeriodO Period Gazetteer, Pleiades (ancient places), Nomisma.org (coins thesaurus). Therefore linking data at least through these vocabularies, which form hubs in the Web of Data, is generally possible.

Some archaeological and other cultural heritage organisations and projects experimented using Linked Data, including ARIADNEplus partners, but thus far a broad take-up in the sector has not occurred. One driver that currently motivates adoption is that more digital cultural heritage projects now turn to the CIDOC-CRM (ontology) for semantic bridging between different collections (Moraitour et al. 2019).

Geser (2016) carried out an extensive review of Linked Data issues. A key issue is the widespread notion of an unfavourable ratio of costs compared to benefits of employing Linked Data standards for information management, publication and integration. The benefits of using Linked Data are quite clear and recent surveys show that librarians and other information professionals are aware of them (McKenna 2018; Smith-Yoshimura 2018).

The benefits include standardisation of own data, interoperability with other publisher's data, reuse of data as well as vocabularies of other providers to enhance own information, linking information across different institutions, greater accuracy and scope in local search results as well as in cross-searching heterogenous data of different providers.

Not as clear are impediments to providing and consuming Linked Data as well as the costs. Regarding impediments McKenna (2018) notes that survey participants faced multiple barriers to using Linked Data particularly in the areas of tooling, data integration, data interlinking, and resource quality. Available tools are often too complex and unsuitable for the needs of heritage institutions. Regarding data integration, for example, respondents indicated that mapping between the different vocabularies used across datasets poses a significant challenge. Respondents had also doubts about the quality and reliability of many currently published Linked Data resources.

The survey results resonate with study results of Geser (2016), who recommended the provision of data mapping recipes and templates, proven tools and guidance material to enable non-IT experts to generate and publish Linked Data. Regarding linkable data he recommended fostering a community of Linked Data curators who ensure the quality and reliability of Linked Data resources.

On the question of costs, the common understanding is that Linked Data are expensive. But so far little research has been published to allow an evaluation and comparison of the costs of current systems, and the advantages of using instead Linked Data methods and technology. Knowing the cost factors and drivers of different Linked Data projects would help reduce the costs, for example by providing dedicated tools, guidance and support for certain tasks.

The main group to convince of the advantages of Linked Data are providers of data repositories and databases for researchers. Geser (2016) noted that take-up of the Linked Data approach by the

[40] Linked Open Data Cloud, https://lod-cloud.net

research community required applications that demonstrate advantages not only regarding data management but also research processes and outcomes.

Activities and results

The task force looked into the current situation of Linked Data and updated the baseline study by Geser (2016) of which some results are described above. The task force also found that other ARIADNEplus partners had trialled using Linked Data, for instance CEIPAC (University of Barcelona) in their Epnet project (completed 2019)[41], and University of Tours (CNRS) in the ongoing OpenArchaeo project (Marlet et al. 2019).

Recently a group of researchers proposed a new Special Interest Group with a focus on Linked Data of Computer Applications & Quantitative Methods in Archaeology (CAA)[42], which may be officially accepted at the next CAA conference. The proponents include members of the ARIADNEplus research community. Thus there is a larger base of experience with Linked Data among ARIADNEplus partners the task force can build upon.

This base can be expanded through existing liaisons with other initiatives, some of which have been active for several years. To give but one example, Pelagios[43], a research community for Linked Data focussed on the ancient world, has remained remarkably dynamic, and is now a network of over 20 members.

There appears to be a generally heightened interest in Linked Data. For instance, the task force lead partner discussed Linked Data issues concerning heritage conservation data and potential synergies with ARIADNEplus with the network project Linked Conservation Data (UK AHRC)[44] and have since joined the second phase of that network, building on the ARIADNEplus work for the Vocabulary Matching Tool. The task lead partner also participated in LODLAM 2020 (Linked Open Data in Libraries, Archives, & Museums) at the Getty Center in Los Angeles and presented the ARIADNEplus vocabulary mapping work at the ITWG (International Terminology Working Group) meeting also at the Getty Center.

Taking account of the current situation regarding Linked Data in the field of archaeology and related fields, and the available knowledge base within ARIADNEplus and of related projects, a working paper has been drafted. This paper outlines how the task force could address requirements for progress in archaeological Linked Data during the lifecycle of the project. The focus will be on the most critical impediments for Linked Data, developing a better understanding of the cost factors, and documenting examples of advantages of Linked Data use in archaeology. A component which plays a major role regarding the advantages, but also the barriers and costs, is the use of vocabularies for semantic data enrichment and linking.

Summary and outlook

A broad take-up of Linked Data methods and technologies for archaeological data has not happened so far. Some archaeological and other cultural heritage organisations and projects have experimented using Linked Data, including ARIADNEplus partners, but for further progress existing impediments must be addressed and removed.

The task force can build on a good base of experience within the project as well as of important related initiatives. These include providers of existing resources in the Linked Data Cloud (Getty Art &

[41] EPnet project, http://www.roman-ep.net/wb/epnet-project/

[42] Data Dragons: SIG on Semantics and LOUD in Archaeology, https://github.com/caa-datadragons

[43] Pelagios Network, https://pelagios.org

[44] Linked Conservation Data, https://www.ligatus.org.uk/lcd/

Architecture Thesaurus, Nomisma.org, PeriodO) and projects such as the Pelagios Network. Previous collaborations with them are being renewed.

An opportunity to update the knowledge base further and forge new relations, will be undertaken during two scheduled round tables at the next CAA conference.[45] The task force will also support the initiative for a new CAA Special Interest Group with a focus on Linked Data.

Future work of the task force will focus on critical impediments for Linked Data, understanding the cost factors more clearly, and providing examples of advantages of using Linked Data in archaeology. Particular attention will be given to advantages but also the costs implied in the use of vocabularies for semantic data enrichment and linking.

4.5 Innovating services with industries

Focus of the task force

The focus of this task force is described as follows (project Task 6.2.4 description): *"A range of industries can benefit from ARIADNEplus' innovative services, including businesses active in heritage preservation, heritage-based regional development, cultural tourism, creative businesses providing heritage presentation services and products. They can access available data and information about sites, monuments and artefacts as well as use services and tools to visualize and communicate objects. Software developers working based on the open source service model may want to provide applications for exploiting data and information for any of the mentioned purposes. The task will identify where and how innovating with certain industries could yield the most significant mutual advantages, explore these opportunities with industry representatives, and provide guidance on how to foster their exploitation."* – Task lead: ICCU, three other assigned partners.

Background

The ARIADNEplus services are being developed primarily for researchers and professionals in the fields of archaeology, heritage management (monuments & sites), and heritage institutions (museums, archives, libraries). In these fields "industry" means contract archaeologists and consultants that are active in so called developer-led or preventive archaeology. They provide fieldwork services for land and urban developers and heritage administrations, mainly excavation work, as for most European countries pre-development evaluation as to whether archaeological remains are present, is not carried out by commercial service providers.

Activity in preventive archaeology by commercial and other organisations in Europe is controlled by governmental institutions and accreditation schemes. The share of commercial service providers in developer-funded excavations varies, e.g. in 2015 in the Czech Republic 15-20% were conducted by 15 licensed private organisations (Mařík 2016: 49), in France around 30% by 19 accredited private firms (Randoin 2016: 62). While in these and other countries most archaeological investigations are conducted by governmental and other public institutions (e.g. museums), in the Netherlands and the UK non-governmental organisations do most of the work. In the UK it is estimated at 90% of all archaeological investigations, in 2015 conducted by 70 members of the Chartered Institute for Archaeologists, of which 46 carried out excavations (Perring 2016: 96-97); however the larger members such as Oxford Archaeology and Wessex Archaeology are charitable trusts.

Professionals active in preventive archaeology have the background to use relevant services and data found through ARIADNEplus, and connected repositories in some countries receive fieldwork reports

[45] The CAA 2020 conference, which was to be held in Oxford, UK, in mid-April has been postponed until later in the year because of the rapid spread of the Corona (COVID-19) virus.

and data from them (e.g. ADS in the UK, DANS in the Netherlands). The task force will explore the interest of archaeological businesses and professionals in using ARIADNEplus services and if there is a need for dedicated services for them.

Two groups which the task force does not consider are commercial repositories and publishers. As Schonfeld & Springer (2019) observed, the request of ever more research funders to make data underpinning published research available through accessible repositories is generating a research data sharing business landscape. They mention examples of commercial repository providers such as Figshare[46] and Mendeley Data[47], and expect more to come. For example, Figshare offers sub-databases for publishers, university libraries (to avoid setting up a data repository), and repository space for research groups. They also bring in data librarians to provide the kind of support services offered by subject-specific databases. ARIADNEplus does not provide a central repository but partner repositories might consider developing special services for publishers and other clients as provided by the commercial repositories.

ARIADNEplus plans to connect digital libraries of archaeological publications and reports, and use the available natural language processing (NLP) service to improve their metadata for indexing and search purposes. NLP of archaeological texts to extract relevant information is a highly specialised undertaking which requires domain knowledge and use of domain-specific vocabularies to be effective (Felicetti et al. 2018; Vlachidis & Tudhope 2016). The improvement of metadata, indexing and search are areas where commercial NLP service providers are active (e.g. NLP of business and web content), but archaeological publications and reports do not represent a niche market for them; from the scientific domains market observers only include healthcare and life sciences as a significant segment (e.g. Markets&Markets 2019).

The task brief mentions businesses active in heritage-based development, cultural tourism, and creative businesses providing heritage presentation services and products. Typical examples of heritage-based development are turning an urban area with historic buildings into a “cultural quarter” or former industrial sites into culture and leisure facilities. The task force does not consider businesses active in this field as relevant users of ARIADNEplus services, nor cultural tourism providers, at least not as direct users of services.

But businesses within the cultural and creative industries seem relevant, particulary if they create products for heritage tourism, sites, monuments and museums. Such products are websites presenting museum collections, multimedia exhibitions, and educational digital games, among others. Therefore the ARIADNEplus services for visual media such as 3D models (and perhaps other services) could be of interest to businesses that are active in this field. For other areas of the cultural and creative industries, e.g. music or audio-visual products, ARIADNEplus services will hardly be relevant (cf. KEA European Affairs 2019).

The task brief also mentions that developers of open source software and services might offer ARIADNEplus applications relevant for research on and/or communication of archaeological sites, monuments and artefacts. The task force will consider this in a later phase of their work.

Activities and results

The task force carried out a first investigation of relevant industries. As the most promising appear to be preventive archaeology and cultural and creative businesses. Among the latter, particularly businesses which create products for heritage tourism, museums and visitor centres of monuments and sites. For the selected businesses the ARIADNEplus services for visual media could be of interest

[46] Figshare (Digital Science), https://figshare.com

[47] Mendeley Data (Elsevier), https://data.mendeley.com

but perhaps others as well. The task force leader and the provider of the ARIADNEplus visual media services, CNR-ISTI's Visual Computing Lab (Pisa), have agreed on a plan for demonstrating the services to relevant businesses.

The plan includes benchmarking of visual media services, identification of the value proposition of the ARIADNEplus services, and development of use cases. Among the use cases, for example, are a multi-layer 3D viewer for excavation layers (preventive archaeology) and 2D-3D viewers for interactive cultural heritage kiosks (cultural & creative businesses). The task force notes that their role is not the marketing of the visual media or other services to businesses. As stated in the task brief, their role is to identify opportunities for innovating services with business users and, where the evaluation shows potential, suggest how to foster exploitation.

Summary and outlook

The task force investigated potential business users of ARIADNEplus services. The group selected preventive archaeology businesses and cultural and creative businesses, which create products for heritage institutions, which as potential users and the visual media services, are particulary relevant. A plan for benchmarking and demonstrating the visual media services to relevant businesses has been developed and is now being implemented, with the goal of highlighting three or four service use cases. In a later phase of the project, examples of innovative uses of ARIADNEplus services and data are foreseen (WP16 Innovative Methods and Pilots). Some of these might also serve as showcases for different services and user groups, beyond archaeological research and heritage management.

5 Contribution to innovation by all project activities

5.1 Overview of the innovation dimensions

The ARIADNEplus innovation strategy aims to promote significant advances in archaeological research and data management in Europe and beyond. The innovation strategy is being supported by all project work packages. The main dimensions of the strategy are:

- *Research policies:* Alignment with the European research policies on FAIR data, Open Science practices, and the European Open Science Cloud (EOSC) initiative.
- *Data integration:* Increase of the ARIADNE data pool, through incorporation of datasets from a greater number of archaeological research domains.
- *Data infrastructure:* Implementation and operation of a Cloud-based platform for data aggregation, integration, discovery, access and use across across institutional and national as well as disciplinary boundaries.
- *Service portfolio:* Provision of enhanced and new services for digital archaeology on the Cloud-based platform.
- *Stakeholder and user base:* Extension of the stakeholder and user base in Europe and beyond, taking account of user needs regarding data, technical services and training.

This chapter describes the contribution of the project work packages to the different dimensions of the innovation strategy, including selected activities and results thusfar.

5.2 Alignment with research policies

The closely related policies of FAIR data and Open Science (see *Section 3.1*) are being promoted by project activities that provide guidance, training and tools for good practice in research data management and sharing. The activities address archaeological research organisations, researchers and repository managers, and are present in different work packages and tasks, e.g WP3 (FAIR Data Management), WP7 (Training task), and WP8-11 (Transnational Access).

Guidelines and support will be provided for Data Management Plans (e.g. a DMP template and wizard for researchers), how to realise FAIRness of data (e.g. data standards, licensing, citation), and data repository setup, management and quality control (e.g. repository creation and CoreTrustSeal accreditation). Training on FAIR data management will be provided to researchers and data managers in the form of short trainings at community conferences, in online webinars and materials, and as part of the Transnational Access offer. TNA is provided by ARIADNEplus centres of expertise in data curation and interoperability (ADS, CNR-ISTI, PIN).

The project management oversees the alignment with the European research policies and takes care that the project follows related ethics requirements (WP1, WP17).

Selected activities and results:

- A roadmap for all ARIADNEplus training and guidance activities has been defined.
- The first round of Transnational Access was organised and some of the selected beneficiaries already visited ARIADNEplus competence centres.
- The first training workshop on FAIR data management, organised by DANS, took place in March 2020, because of the COVID-19 emergency situation in a reduced, online format.

5.3 Data integration

ARIADNEplus will increase the pool of integrated research data through incorporation of datasets from more archaeological research communities, e.g. paleo-anthropology, bio-archaeology, inorganic materials study, environmental archaeology, maritime and underwater archaeology, remote sensing, among others. This requires definition of application profiles for the different datasets, support in mapping to the ARIADNEplus Data Model, and standardised description of the data records with appropriate general and domain-specific vocabularies. Tools are being provided by project partners, e.g. a tool for mapping databases to the ARIADNEplus Data Model (FORTH's 3M Tool)[48], and a tool for mapping "local" to common domain vocabularies (USW's Vocabulary Matching Tool)[49]. Where support in the design of national or thematic data infrastructures is needed, ARIADNEplus partners will provide this in view of incorporating their data.

These activities are carried out in WP4, WP5 and WP14 with a focus on standardising data description, and increasing the depth of data integration (item-level access). Updated existing ARIADNE datasets and datasets of new reseach communities will be ingested and integrated in the new data catalogue. A dashboard will allow monitoring of the knowledge integration achieved in the data catalogue and portal.

Selected activities and results:

- A richer ARIADNEplus Data Model was defined, building on the data catalogue model for the humanities developed in the Parthenos cluster project.
- The development of application profiles for most of the different types of domain data is underway, also taking into account data from new interested projects (e.g. ROCEEH - The Role of Culture in Early Expansions of Humans, in a joint workshop in January 2020).
- The tools for mapping databases to the ARIADNEplus Data Model and describing data records (metadata) with common domain vocabularies are already ported to the Cloud-based platform.
- A manual and helpdesk to support data providers are available.

5.4 Data infrastructure

An overhaul of the ARIADNE data infrastructure is necessary to allow effective incorporation of the many new datasets from different archaeological domains and provision of advanced and new services on a Cloud-based platform (D4Science). The setup and operation of the new data infra-structure is to be undertaken within WP12 and WP13. The work comprises of the implementation of the new data aggregation infrastructure, data catalogue and portal, underlying knowledge base, and knowledge discovery and browsing engine. Furthermore the foreseen virtual research environments (VREs) will be designed, implemented and operated. The technical work packages will also take care of the future alignment of the ARIADNEplus catalogue of services and datasets with the respective catalogues of the European Open Science Cloud (for more information on the alignment see *Section 3.1.4*).

Selected activities and results:

- The new data aggregation infrastructure and data catalogue are implemented.
- Aggregation of updated and new datasets is being tested.

[48] Mapping Memory Manager - 3M tool (FORTH-ICS), https://isl.ics.forth.gr/3M/

[49] Vocabulary Matching Tool (USW), http://heritagedata.org/vocabularyMatchingTool/

5.5 Portfolio of services

ARIADNEplus will extend the service portfolio with new and enhanced services on the Cloud-based D4Science platform. The goal is to allow innovative and effective ways of carrying out data-based research in archaeology (digital archaeology).

This will be enabled by the creation of different ways to query the ARIADNEplus knowledge and database, services for visual media (e.g. 3D models), services for annotation of texts and images (e.g. fieldwork reports, artefact or laboratory images), text mining and natural language processing services, and geo-spatial/GIS data services; an overview of the services is given in *Section 7.2* (Annex B). The services will be implemented in WP15, while a number of pilots in WP16 will demonstrate how ARIADNEplus services and data can be used in innovative ways.

Selected activities and results:

- ARIADNEplus services for end-users are being ported on the Cloud-based platform.
- Already available are the Visual Media services of CNR-ISTI's Visual Computing Lab and the Natural Language Processing service developed by USW.

5.6 Extension of the stakeholder and user base

The extension of the ARIADNEplus stakeholder and user base, both for data providers and end-users, is being supported by WP2. Some communities have already been involved in ARIADNE, such as researchers and professionals in sites & monuments, surveys, archaeological artefacts. Now further communities are being mobilised and engaged, including researchers in palaeoanthropology, bio-archaeology, environmental archaeology, maritime and underwater archaeology, material sciences and dating, among others.

Special attention is being devoted to coordination with major archaeological agencies and associations (e.g. European Archaeological Council, European Association of Archaeologists, CAA International), new institutional partners in Central and Southeastern Europe, and international collaboration with non-European institutions and networks (e.g. in Japan, Argentina and USA).

WP2 also includes the task of surveying user needs and ensuring their demands are being met by the ARIADNEplus technical services, and with other services, such as training. Provision of the services in line with the needs of the user base will be crucial to achieve the expected innovation and impact in the archaeological sector in Europe and beyond.

The project dissemination, and involvement of stakeholders and users, are being supported by WP7, through the project website, social media channels, newsletters and other publications, as well as help with organising training (e.g. webinars) and presentations at domain conferences.

Selected activities and results:

- The ARIADNEplus communication channels are set up and regularly fed with information on project activities and opportunities to participate.
- "The ARIADNE Impact" book, published in October 2019, presents case studies of how the ARIADNE initiative has shaped many of the current data management practices in archaeology and will continue to do so.
- The first round of reviewing the community needs was carried out based on a large online survey (summarised results are presented in this report).
- A number of conference sessions and workshops at international and national events have been held and future events organised.

6 From innovation to impact

The ARIADNEplus innovation strategy aims to bring about significant advances in archaeological research and data management, within different areas of innovation and impact. The sections of this chapter present an overview of the ARIADNEplus impact indicators, describe their mapping to the impacts expected by the Horizon 2020 Work Programme, and give an outlook on the further work and deliverables on project innovation and impact.

6.1 Overview of the ARIADNEplus impact indicators

For the different areas of ARIADNEplus innovation and impact, a set of indicators of success has been defined and mapped to the impacts expected of Integrating Activities for Advanced Communities, funded under the Horizon 2020 Work Programme 2018-2020, for European Research Infrastructures (including e-Infrastructures).[50] After the successful ARIADNE project, ARIADNEplus is now such an Integrating Activity.

The set of ARIADNEplus impact indicators is presented in the tabular overview below. The set is comprised of indicators that were already defined in the Description of Work (Part B: Impact) for the project, and others that were added to cover contributions by all project activities. The tabular overview includes the mapping of the ARIADNEplus indicators to the impacts expected by the Work Programme.

The programme states 10 expected impacts, and the mapping is indicated by a code, [EI-1] to [EI-n], which refers to the sequence of the statements, however some of the statements require a significant amount of background information to be considered appropriately. Expectations may not apply to a certain type of research infrastructure in general, or need to be specified for that type. For example, the statements do not distinguish between different types of infrastructures, e.g. between a large natural science laboratory and a digital infrastructure for research such as ARIADNEplus. Furthermore, the expectations have not been updated to take account of the new research policies such as FAIR data, Open Science practices, and the European Open Science Cloud.

Therefore the expectations had to be interpreted and specified so that they could be applied as indicators of success of ARIADNEplus. The background and coverage of the indicators is described in the next section.

[50] Horizon 2020 Work Programme 2018-2020. 4. European research infrastructures (including e-Infrastructures), INFRAIA-01-2018-2019: Integrating Activities for Advanced Communities, p.53, https://ec.europa.eu/research/participants/data/ref/h2020/wp/2018-2020/main/h2020-wp1820-infrastructures_en.pdf

Table 1: Overview of the ARIADNEplus impact indicators.

Extension of the stakeholder and user community [EI-4], [EI-5]
Archaeological and cultural heritage institutions involved (total number of institution)
European institutions
Non-European institutions
Collaboration with archaeological agencies and associations (EAC, EAA, CAA,…)
Partnership with industry [EI-4]
Businesses of contract archaeologists in developer-led/preventive archaeology
Cultural and creative businesses with a focus on cultural heritage institutions
Researchers and data managers [EI-5]
Researchers and data managers involved (total number, not incl. use of services)
Participants in the community needs survey and expert panels
Participants in presentations at domain events
Participants in Transnational Access (TNA)
Participants in FAIR data and other training (e.g. tutorials, short courses, workshops)
Users of the A+ Data Management Plan template and wizards
Research infrastructures coordination & standardisation [EI-3]
Coordination with other e-infrastructures (humanities, heritage and other)
European-level thematic e-infrastructures
National-level thematic e-infrastructures
E-infrastructures standardisation
Harmonisation of data infrastructures (data catalogues, services)
Collaboration with domain standards groups (vocabularies)
Mapping of A+ and EOSC catalogue models

Data providers and datasets [EI-8]
Data providers from different countries (total)
European data providers
Non-European data providers
New repository initiatives supported; repositories accredited (CoreTrustSeal)
Data integration and management
Datasets from different archaeological research domains in the data portal (total of data records)
Dataset application profiles created and mapped to the A+ data model
Local data vocabularies mapped to common vocabularies
New data vocabularies provided and applied
Datasets integrated (updated, newly ingested)
Datasets integrated at item level
Archaeological/CH digital libraries integrated
Research e-infrastructure services & users [EI-1], [EI-2]
Provision of Cloud-based data and research services
Users of the data portal (data discovery and access)
Users of reseach services (visual media, annotation, NLP, GIS,…)
Users of Virtual Research Environments (combinations of services) for research communities
Demonstrators of innovative uses of services and data
Cross-disciplinary fertilisations and sharing of resources [EI-6]
Researchers and developers from many archaeology and heritage domains
Use of common data models and vocabularies (e.g. CIDOC-CRM extensions)
Collaboration with external Linked Data adopters

A+ resources (services, datasets) in the EOSC catalogues	Sharing of open source software tools
Dissemination of information and knowledge	Sharing knowledge: presentations, papers, guidelines (see: Dissemination)
Project website (unique visits)	**Use of services beyond research** [EI-9]
Social media/network members and reach	User groups beyond archaeological research (estimated total)
(Co-)organised events (stakeholder meetings, conference sessions, workshops)	Citizen scientists (e.g. archaeological finds by amateurs, local heritage projects)
Distribution of newsletters, policy briefs, guidelines, training material, etc.	Educators and students (e.g. study work)
Papers in journals, conference proceedings, book chapters	Preventive archaeology businesses, cultural & creative businesses (see: Partnership with industry)

6.2 Mapping to the RI programme expected impacts

The mapping of the ARIADNEplus impact indicators to the impacts expected by the H2020 Work Programme for Research Infrastructures is presented as follows: The expected impacts are addressed according to their sequence (indicated by [EI-1] to [EI-n]) and placed under the relevant headings of the tabular overview. Expected impacts which are not relevant for ARIADNEplus are addressed in the final section.

Two parts of the tabular overview, although essential, are not described in dedicated sections because these activities are not emphasised in the expected impacts of the Work Programme. These are the extension of the stakeholder community, e.g. archaeology and heritage agencies, institutions and associations involved, and dissemination activities such as conference sessions, scientific papers, distribution of policy briefs, guidelines, etc.

Research e-infrastructure services & users

Here two expectations of the RI programme are addressed:

[EI-1] *"Researchers will have wider, simplified, and more efficient access to the best research infrastructures they require to conduct their research, irrespective of location. They benefit from an increased focus on user needs."*

[EI-2] *"New or more advanced research infrastructure services, enabling leading-edge or multi-disciplinary research, are made available to a wider user community."*

The ARIADNEplus e-infrastructure and services integrate data infrastructures (data repositories, databases) from across Europe and beyond, and enable efficient search, access and (re)use of data irrespective of the location where researchers conduct their research. The development of ARIADNEplus takes account of the needs identified by the user community regarding data sharing, access and (re)use, new services, and related training (see *Section 3.3*).

Archaeology is a multi-disciplinary field of research in which researchers require knowledge and data from different domains of research. Therefore, ARIADNEplus integrates data resources from far more archaeological domains than the initial ARIADNE project. New and advanced services are being implemented on a Cloud-based platform (D4Science) to accommodate this expansion. Based on the Virtual Research Environment (VRE) approach, different services can be combined and configured to provide a VRE for the specific needs of a domain of research. The project will also demonstrate

advantages of using ARIADNEplus services and data to the archaeological research community and other user groups.

The ARIADNEplus impact indicators here highlight the usage of the Cloud-based services, e.g. users of the data portal, research services and VREs, and the demonstrators of innovative uses of services and data.

Research infrastructures coordination & standardisation

[EI-3] *"Operators of related infrastructures develop synergies and complementary capabilities, leading to improved and harmonised services. There is less duplication of services, leading to an improved use of resources across Europe. Economies of scale and saving of resources are also realised due to common development and the optimisation of operations."*

ARIADNEplus coordinates its activities with related e-infrastructures in the field of humanities and heritage research at the European and national levels, e.g. CLARIN, DARIAH, E-RIHS, Europeana Research, Huma-Num (France), ICDI (Italy), and others. The goal here is to develop synergies and exploit complementary capabilities. Joint development in ARIADNE will be continued regarding harmonised data models and vocabulary, e.g. data catalogues, use of the CIDOC CRM and domain extensions of the ontology, and appropriate domain vocabularies.

The overall strategy of ARIADNEplus regarding improved use of resources, economies of scale and cost-savings is Cloud-based virtualisation and integration in the European Open Science Cloud (EOSC). This will allow other research e-infrastructures in the EOSC eco-system using complementary ARIADNEplus resources (services, data) relevant for their operation, and vice versa.

The ARIADNEplus impact indicators cover the coordination with other e-infrastructures (humanities, heritage and other) both at the European and national level, the collaboration on common standards, and the integration of ARIADNEplus resources in the EOSC catalogues (services, datasets).

Partnership with industry

[EI-4] *"Innovation is fostered through a reinforced partnership of research organisations with industry."*

The notion of industrial innovation in the Work Programme is informed by research infrastructures (RIs) in fields such as energy, materials or life sciences research and development. In these fields innovation may result from advances in the construction of RIs involving industrial providers of components (e.g. procurement of innovative instrumentation), joint R&D projects, and the use of experimental facilities by industrial researchers and developers (ESFRI Innovation Working Group 2018).

The evaluations of the RI Programmes found that industrial businesses play no leading role, if any, in the building of e-infrastructures within Integrating Activities (EPIRIA 2014: 60; Chang et al. 2017: 21). The e-infrastructures are not developed by industrial actors, but by domain and technological research organisations, and the ARIADNE initiative is an exemplary case of such collaboration.

In archaeology, commercial actors are businesses employing contract archaeologists, and are active in so called developer-led or preventive archaeology. They provide professional fieldwork services for land or urban developers and heritage administrations. In some European countries commercial contract archaeology has become the dominant form of field archaeology, while in others (semi-) public bodies play a greater role (see *Section 4.5*). These can benefit from using ARIADNEplus services and datasets, but may also share fieldwork reports and data through underlying repositories.

Access to ARIADNEplus services and data may also be of interest to businesses of the cultural and creative industries, particulary if they create products for heritage tourism, sites, monuments and

museums. Such products are websites presenting museum collections, virtual reconstructions, multimedia exhibitions, and educational digital games, among others. Therefore the ARIADNEplus services for visual media, such as 3D models, are considered relevant for businesses active in this field. The ARIADNEplus task force "Innovating services with industries" addresses such businesses to explore and demonstrate opportunities (see *Section 4.5*).

The ARIADNEplus impact indicators here cover collaboration with preventive archaeology businesses and relevant businesses of the cultural and creative industry.

Researchers and data managers involved

[EI-5] *"A new generation of researchers is educated that is ready to optimally exploit all the essential tools for their research."*

Researchers need tools for next-generation digital research in archaeology and cultural heritage, and at the same time, are asked to acquire skills in data management and sharing based on the FAIR data principles. ARIADNEplus supports these requirements through provision of advanced services, Transnational Access to partner competence centres, FAIR data and other training (e.g. tutorials, short courses, workshops), a Data Management Plan template and wizards for archaeological data. Furthermore, guidelines on FAIR good practices will be provided for researchers and data managers.

The impacts expected by the Work Programme do not take account of the new research policies regarding FAIR data and Open Science, which concern practices related to data and publications generated based on using research tools. The project impact indicators highlight participation of researchers in activities focused on FAIR data management and sharing (e.g. Transnational Access and other training). In addition, participation in the community needs survey and expert panels, as well as in sessions and workshops at domain events, are included. Use of ARIADNEplus data and research services is addressed under "Research e-infrastructure services & users".

Cross-disciplinary fertilisations and sharing of resources

[EI-6] *"Closer interactions between larger number of researchers active in and around a number of infrastructures facilitate cross-disciplinary fertilisations and a wider sharing of information, knowledge and technologies across fields and between academia and industry."*

The participation of research groups and mobilisation of datasets from many archaeology and heritage research domains within ARIADNEplus will foster cross-disciplinary fertilisations and a wider sharing and take-up of knowledge and technologies for data sharing, integration and (re)use. Cross-fertilisation is generally fostered by the common objective of research organisations and groups to integrate data and enable comparative research and synthesis. This requires use of appropriate data models, ontologies (e.g. CIDOC CRM extensions), and domain vocabularies, which allow interlinking and exploitation of shared data based on Linked Data methods and technologies.

The ARIADNEplus new and advanced services will be available for users from all relevant fields, including archaeological research organisations, businesses active in preventive archaeology, heritage administration/management, cultural and creative businesses providing services for museums or heritage tourism organisation, and others. Developers of open source software and services may offer ARIADNEplus additional technical applications for research on and communication about archaeological sites, monuments and artefacts.

The related ARIADNEplus impact indicators cover the participation of researchers and developers from many archaeology and heritage domains, use of common data models and vocabularies, collaboration with external Linked Data adopters, and sharing of open source software tools. In addition the sharing of information and knowledge is covered by indicators for dissemination

activities, e.g. (co-)organised events such as conference sessions and workshops, research publications, distribution of guidelines, etc.

Data integration and management

[EI-8] *"The integration of major scientific equipment or sets of instruments and of knowledge-based resources (collections, archives, structured scientific information, data infrastructures, etc.) leads to a better management of the continuous flow of data collected or produced by these facilities and resources."*

The integration and management of knowledge-based data resources comes late on the list of impacts expected by the Work Programme, but is a core focus of Integrating Activities regarding e-infrastructure. The e-infrastructure is put in place to allow aggregation, integration, discovery and access of data from distributed digital archives and databases, i.e. the flow from data generators to data users. In most cases, including the ARIADNEplus e-infrastructure, there is no direct access to the raw data generated with scientific instruments but data prepared for (re)use in connected digital repositories and databases.

Through the ARIADNEplus e-infrastructure, researchers and other users will be able to discover and access data resources held by integrated repositories and databases, from around Europe and non-European countries (Argentina, Israel, Japan, USA).

Related impact indicators include the number of data providers and datasets from different archaeological research domains within the data portal, dataset application profiles created and mapped to the common data model (which can be used also by new data providers), use of common vocabularies for data description, and datasets integrated at item level, among others.

Use of services beyond research

This is another sub-heading of the indicators for research e-infrastructure services and users, which addresses the following RI programme expectation:

[EI-9] *"When applicable, the integrated and harmonised access to resources at European level can facilitate the use beyond research and contribute to evidence-based policy making."*

Serving many users groups beyond archaeological researchers and data managers (repositories, databases) is not a core objective of ARIADNEplus. The main user groups include also heritage administrations/managers and professionals in preventive archaeology, while others will be addressed where appropriate.

Potential other users of the ARIADNEplus data portal and services are citizen scientists, educators and students. For example, the project integrates metal-detector and other finds by amateurs in countries where this is permitted, and registration of finds in national databases has been enabled. Science educators and students could benefit from the availability of natural sciences data on archaeological remains. Access to relevant content may also stimulate study work and visits to archaeological museums and sites.

Policy-making for archaeological and cultural heritage is a complex matter, due to conflicting interests, especially protection vs. exploitation, different implementation of international conventions and actual heritage management practices across countries (Europae Archaeologiae Consilium 2018). Contributions to evidence-based policy making by researchers who use ARIADNE-plus services and data may be possible, but this has not been defined as a core focus of the ARIADNE initiative in its current development phase.

ARIADNEplus impact indicators for use of services beyond research mainly consider "citizen scientists" (e.g. archaeological finds by amateurs, local heritage projects), and educators and students (e.g. study work); business users are covered under "Partnership with industry" (preventive archaeology and cultural & creative businesses).

Not applicable expectations

[EI-7] *"For communities which have received three or more grants in the past, the sustainability of the integrated research infrastructure services they provide at European level is improved."*

[EI-10] *"When applicable, the socio-economic impact of past investments in research infrastructures from the European Structural and Investment Funds is enhanced."*

The archaeological research community so far has been supported by two Integrating Activity grants (ARIADNE and ARIADNEplus), nevertheless, the initiative strives to sustain the implemented research infrastructure services at the European level beyond the current project lifecyle.

The expectation of socio-economic impacts concerns large physical research infrastructures (e.g. major natural or life sciences laboratories) built with investments from the European Structural and Investment Funds. These research infrastructures (RIs) should demonstrate significant socio-economic impacts in the region, e.g. generation of employment, economic prosperity and growth. The ARIADNE initiative does not focus on such RIs and indicators, as their socio-economic impacts cannot be easily applied to research e-infrastructures (virtual RIs), for which indicators such as improved access to data and research services are used.

6.3 Outlook on further work and deliverables

Monitoring and evaluation

The impact monitoring task (T6.3) will further detail the set of impact indicators and include appropriate target values. Some of the values are defined within the Description of Work (Part B: Impact) but many additional indicators need to elaborated upon and agreed to cover all project activities with observable quantitative or qualitative target values. The monitoring task will also provide templates, guidance and support for documenting the results of the different activities, including use of the technical services, training, and knowledge transfer, among others.

A first overall impact evaluation report will be delivered in project month 30 (June 2021). Bringing together and analysing the monitored results will allow adjustments and/or fine-tuning of activities and of their interplay, with the goal of optimising project outcomes and impacts. The full account and evaluation of the outcomes and impacts will be presented in project month 48 (December 2022).

Final innovation and sustainability plans

The innovation strategy is a "living document" that will be updated based on new knowledge and opportunities, such as newly formed expert networks, initiatives within research and data communities, and advanced technologies, among others. Special attention will be paid to the development of the European Open Science Cloud, its setup, rules of engagement, and emerging "marketplace" of services and data. Furthermore, the innovation strategy will take account further insights regarding needs and requirements of the ARIADNEplus user community.

The final innovation strategy will be delivered in project month 40 (April 2022). The deliverable will report the results of innovation-focused activities carried out, and present the programme of further planned activities, within and beyond the formal lifecycle of the project.

The ARIADNEplus sustainability plan, a business model for the service provision beyond the funded period of the project, will be developed by the project coordinator and the steering committee (i.e. all WP leaders). The final plan will be presented in project month 48 (December 2022), including projections of user demand, cost factors, and required resources to maintain and further develop the portfolio of services. It will also report the expressed interest of stakeholders to support the proposed sustainability measures and business model.

7 Annexes

7.1 Annex A: Overview of new thematic areas of data resources

The first ARIADNE project e-infrastructure primarily integrated monument and site inventories, excavation archives, fieldwork reports, and artefact databases. ARIADNEplus aims to incorporate a wider range of data types from different archaeological domains of research. The new thematic domains include:

- Human Palaeo-biology and Palaeo-environments
  - Palaeo-anthropology
  - Bio-archaeology and Ancient DNA
  - Environmental Archaeology
- Analytical Investigations
  - Inorganic Materials Study
  - Dating, including Dendrochronology
- Archaeological Prospection
  - Field Survey
  - Metal Detector Survey (where allowed by the law) and archaeological finds
  - Remote Sensing (e.g. LiDAR and satellite data)
- Monuments and Sites
  - Spatio-temporal Data (GIS)
  - Archaeological Fieldwork
  - Burial Archaelogy
  - Standing Structures
  - Maritime and Underwater Archaeology
- Inscriptions
  - This category comprises proper inscriptions, as analysed by epigraphy, inscriptions made on objects (e.g. on amphoras), and rock carvings.

Moreover, ARIADNEplus will integrate archaeological publications from digital libraries. The project will exploit major sources such as OpenAIRE[51] and individual journals such as Internet Archaeology[52], Archeologia e Calcolatori[53], and others.

[51] OpenAIRE, https://www.openaire.eu

[52] Internet Archaeology (journal), https://intarch.ac.uk

[53] Archeologia e Calcolatori (journal), http://www.archcalc.cnr.it

7.2 Annex B: Overview of services for researchers and data managers

This overview presents the planned ARIADNEplus services for end-users, researchers (archaeologists, laboratory-based scientists and others) as well as data managers (repositories, databases). The overview includes new as well as existing but improved services. Not included are "back-office" services, i.e. those which run the service provision platform and others which end-users do not use directly.

Existing services

The following services are already available on the data portal and therefore not included in the list of new or advanced innovative services for end-users (ARIADNEplus WP15):

- Register a dataset in a portal that allows searching data from many providers
- Discover and access archaeological data stored in repositories in different European and other countries
- Multi-lingual search for archaeological data
- Spatially and/or chronologically defined search options

Some improvements of these services are possible and planned, or already prepared (e.g. enrichment of the catalogue model of the dataset registry). Regarding the options for data search ARIADNEplus aims to implement semantic search based on metadata and vocabularies in Linked Data formats.

Enhanced and new services

Search and visualise geo-spatial / GIS datasets

These services will be supported by the D4Science Cloud-based Geoserver, which already has many of the required functionalities built-in. The GIS services comprise the usual services present in GIS systems, for example buffer definition, layer selection, proximity, viewshed analysis and so on. The services and tools of the Geoserver will allow integration of archaeological geo-information provided by partners, and support the GIS functionalities of the space-time services (e.g. improvement to the exisiting spatially and/or chronologically defined searches). In order to allow this the Geoserver will also rely on gazetteers and named time period vocabularies developed within the project.

Visual content services

Services for visualisation and manipulation of archaeological imagery developed by the CNR-ISTI Visual Computing Laboratory have already been a well developed part of the ARIADNE service portfolio. These and new visual content services will be enhanced in ARIADNEplus building on advances in recent projects of partners such as VisualMedia EOSCpilot Science Demonstrators[54].

Display and manipulate visual data objects (e.g. RTI images, 3D models, LiDAR data)

These services allow fast and efficient online rendering and manipulation of advanced forms of visual content such as Reflectance Transformation Imaging (RTI), 3D models and LiDAR imagery. The services are already Cloud-based, but will be adapted to the specific needs of archaeologists, and integrated into the ARIADNEplus service interface (in development).

Link and present together visual media (e.g. a 3D model) and related documentation

This will be a toolkit allowing effective organisation of visual content from archaeological research (e.g. 3D models of artefacts or monuments), including linking with other research documentation,

[54] EOSCpilot: Science Demonstrators, http://eoscpilot.eu/science-demonstrators

and visualising how they relate to each other. The toolkit comprises tools developed by the CNR-ISTI Visual Computing Laboratory. Apart from some foreseen improvements, the toolkit will be ported to the ARIADNEplus Cloud environment and adapted to the service interface.

Visualise in 3D the layers of an excavation and the related documentation

This service will allow visual documentation of archaeological excavations with a focus on 3D documentation, specifically 3D visualisation of the excavation layers and related documentation. It will build on the Ephemera service developed by ARIADNEplus partner CYI-STARC[55]. As with other services porting to the ARIADNEplus Cloud environment, an adaption to the service interface is required.

Annotate images (e.g. artefact or laboratory images) and link them with other content

These services will allow annotation and linking of research images with other content such as protocols, documentation or published articles. The services will support semi-automatic annotation and linking of a smaller number of content items, using relevant vocabularies (e.g. for cultural content such as inscriptions). One specialised tool that is foreseen is the DAP tool[56] for archaeological objects, which contain written/symbolic information. Among other functionalities it allows CIDOC-CRM compliant annotation of images of such objects. The tool was developed by the ARIADNEplus partners Archaeological Museum of Zagreb (AMZ) and CNR-ISTI Visual Computing Laboratory. Porting of this tool to the ARIADNEplus Cloud environment and adaption to the service interface is required.

Textual content services

Annotate texts (e.g. fieldwork or laboratory reports) and link them with other content

These services will support semi-automatic annotation and linking of a smaller number of documents, such as archaeological reports and other content. One candidate for developing such services is the open source annotation system Pundit[57], which has been used by several digital humanities projects. In ARIADNEplus the services will need to support using archaeological and other scientific vocabularies (e.g. scientific work such as material analyses).

Identify & extract information from textual sources (e.g. a document repository) to produce metadata

This service is based on the previous ARIADNE text mining and NLP tool, further developed into TEXTCROWD, a Cloud-based NLP tool created as a Science demonstrator in the EOSCpilot EU project (Felicetti et al. 2018).[58] Further development will include extension regarding the NLP functionality and languages (so far English, Italian, Dutch), and porting on the ARIADNEplus Cloud environment, as well as adaptation to the service interface.

Data vocabularies, mapping and linking

The following three services are mainly intended for data managers (repositories, databases) to enhance and employ vocabularies with the goal of linking and integrating their own and other datasets, based on Linked Data standards and technologies.

Align own vocabulary terms with international thesauri (e.g. Getty Art & Architecture Thesaurus)

The Vocabulary Matching Tool developed by the Hypermedia Research Group of ARIADNEplus partner University of South Wales allows aligning of a vocabulary (term list, thesaurus) with common

[55] Ephemera (CYI-STARC), http://ephemera.cyi.ac.cy

[56] Digital Autoptic Process (DAP) tool, http://tss.isti.cnr.it/dap

[57] Pundit, http://thepund.it

[58] EOSCpilot: Science Demonstrators, http://eoscpilot.eu/science-demonstrators

vocabularies such as the Getty Arts & Architecture Thesaurus. The tool is already available as a service in the ARIADNEplus Cloud environment.

Map a database (schema) to the CIDOC-CRM extended for archaeological research data

The Mapping Memory Manager (3M) system developed by ARIADNEplus partner FORTH-ICS allows the mapping of database schema to the CIDOC-CRM (ontology), including the extensions developed in ARIADNE for archaeological research data (e.g. excavation, standing structures, epigraphy).

8 References

Abdulrahman, Ammar (2017): The Damage Done: The Arab Spring, Cultural Heritage and At-Risk Archaeologists. Arab-German Young Academy of Sciences and Humanities (AGYA) – Transformation Group Working Paper No. 1/2017, available at http://agya.info/working-groups/transformation/working-paper-series-academia-in-transformation/

Agabirwe, Patience (2018): The influence of the open access movement in faculty research output and university libraries in Africa, pp. 227-250, in: XXIII SCECSAL, Entebbe, Uganda, 23-27 April 2018, https://www.scecsal.org/publications/papers2018/013_agabirwe_2018.pdf

ALLEA (2020): Sustainable and FAIR Data Sharing in the Humanities: Recommendations of the ALLEA Working Group E-Humanities. Edited by N. Harrower, M. Maryl, B. Immenhauser & T. Biro. ALLEA - All European Academies, Berlin, February 2020, https://doi.org/10.7486/DRI.tq582c863

Aloia N., Binding C., Cuy S., Doerr M., et al. (2017): Enabling European Archaeological Research: The ARIADNE E-Infrastructure. In: Internet Archaeology 43, https://doi.org/10.11141/ia.43.11

Archaeology Data Service (University of York, UK), http://archaeologydataservice.ac.uk

Archeologia e Calcolatori (journal), http://www.archcalc.cnr.it

Arches (open source data management platform for the heritage field), https://www.archesproject.org

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

ARIADNE (2017a): Final Impact Evaluation Report, D2.5, January 2017. Available http://legacy.ariadne-infrastructure.eu/resources-2/deliverables/

ARIADNE (2017b): Report on E-Archaeology Frameworks and Experiments, D17.1, January 2017. Available http://legacy.ariadne-infrastructure.eu/resources-2/deliverables/

ARIADNEplus (2019): Initial Report on Community Needs. D2.1, 31 October 2019, https://ariadne-infrastructure.eu/wp-content/uploads/2019/11/ARIADNEplus_D2.1_Initial-Report-on-Community-Needs-1.pdf

Asogwa, Brendan E. (2011): Digitization of Archival Collections In Africa For Scholarly Communication: Issues, Strategies, and Challenges. In: Library Philosophy and Practice (e-journal), 651, https://digitalcommons.unl.edu/libphilprac/651

A-Z - The Archives Network Israel, http://www.a-z.digital/en/about/

Beck A. & Neylon C. (2012): A vision for Open Archaeology. In: World Archaeology, 44(4): 479-497, http://www.tandfonline.com/toc/rwar20/44/4#.VEPIOskucyg

Beck, Anthony (2013): Opening access to heritage resources: risk, opportunity or paradigm shift?, pp. 40-47, in: Archäologie und Informationssysteme, Arbeitshefte 42, http://www.landesarchaeologen.de/fileadmin/Dokumente/News/NI_Archaeologie_und_Informationssysteme_Arbeitsheft42.pdf#page=4

Bibliotheca Alexandrina, https://www.bibalex.org

Bibliothèque Numérique Marocaine, http://www.bnrm.ma/bnrm/fr/bibliotheque/nos-collections.html

CAA - Computer Applications & Quantitative Methods in Archaeology, https://caa-international.org

Casana, Jesse (2013): Archaeology after the Arab Spring. In: ASOR - The Ancient Near East Today, Vol. I, No. 2, May 2013, http://www.asor.org/anetoday/2013/05/archaeology-after-the-arab-spring/

Chang H., Game A., Profumo F., Semeria M.N. & Žic-Fuchs M. (2017): Interim Evaluation of the European Research Infrastructures including e-Infrastructures in Horizon 2020. Report of the Expert Group. European Commission, DG Research and Innovation, September 2017, https://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/ri_interim_evaluation_expert_report_112017.pdf

CIDOC Conceptual Reference Model, http://www.cidoc-crm.org

Coalition for Archaeological Synthesis (CfAS), USA, http://www.archsynth.org

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

Cox S. & Yu J. (2018): Oznome - Data ratings (22 May 2018), https://confluence.csiro.au/display/OZNOME/Data+ratings

CULTNAT - Center for Documentation of Cultural and Natural Heritage, http://www.cultnat.org

D4Science (Institute of Information Science and Technologies, ISTI-CNR, Pisa, Italy), https://www.d4science.org

Data Dragons: SIG on Semantics and LOUD in Archaeology, https://github.com/caa-datadragons

di Lernia S. & Gallinaro M. (2019): Introducing the 1st Workshop on “Archaeology in Africa, pp. 1-5, in: Archaeology in Africa. Potentials and perspectives on laboratory & fieldwork research. Arid Zone Archaeology, Monographs Series 8, Sapienza Università di Roma, Italy, https://www.academia.edu/40844865/

di Lernia, Savino (2015): Save Libyan archaeology. In: Nature, Vol. 517, Issue 7536, 547–549, https://www.nature.com/news/cultural-heritage-save-libyan-archaeology-1.16781

Digital Autoptic Process (DAP) tool (CNR-ISTI Visual Computing Laboratory) http://tss.isti.cnr.it/dap

Digital Repository of Ireland, https://www.dri.ie

DiSSCo - Distributed System of Scientific Collections (ESFRI Roadmap project), https://www.dissco.eu

EAMENA - Endangered Archaeology of the Middle East and North Africa, https://eamena.arch.ox.ac.uk

E-Depot for Dutch Archaeology (DANS, Netherlands), http://www.edna.nl

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 (2012): e-IRG Roadmap 2012, http://e-irg.eu/documents/10920/12353/e-irg_roadmap_2012-final.pdf

e-IRG - e-Infrastructures Reflection Group (2013): White Paper 2013, http://e-irg.eu/catalogue/eirg-1003

e-IRG - e-Infrastructures Reflection Group (2016): Roadmap 2016, http://e-irg.eu/documents/10920/12353/Roadmap+2016.pdf

e-IRG - e-Infrastructures Reflection Group (2019): National Nodes - Getting organised; how far are we? Implementing e-Infrastructure Commons and the European Open Science Cloud. e-IRG, June 2019, http://e-irg.eu/catalogue/eirg-1006

EOSC Executive Board (2019): European Open Science Cloud (EOSC) Strategic Implementation Plan. European Commission, DG Research and Innovation, Brussels, June 2019, http://dx.doi.org/10.2777/202370

EOSC Portal, https://www.eosc-portal.eu

EOSC-Hub (H2020, 1/2018-12/2020), https://www.eosc-hub.eu

EOSC-Pillar (H2020, 7/2019-6/2022), https://www.eosc-pillar.eu

EOSCpilot - Science Demonstrators (Textcrowd, VisualMedia), http://eoscpilot.eu/science-demonstrators

Ephemera (CYI-STARC, Cyprus), http://ephemera.cyi.ac.cy

EPnet - Production and Distribution of Food during the Roman Empire: Economic and Political Dynamics (ERC Advanced Grant project, CEIPAC, University of Barcelona), http://www.roman-ep.net/wb/epnet-project/

E-RIHS - European Research Infrastructure for Heritage Science (ESFRI Roadmap project), http://www.e-rihs.eu

ESFRI - European Strategy Forum on Research Infrastructures, Innovation Working Group (2018): Innovation-oriented cooperation of Research Infrastructures. ESFRI Scripta Volume III, https://www.esfri.eu/sites/default/files/ESFRI_SCRIPTA_VOL3_INNO_single_page.pdf

Europae Archaeologiae Consilium (2018): Making Choices: Valletta, Development, Archaeology and Society. A Report of the 'Making Choices' Working Group of the Europae Archaeologiae Consilium, March 2018, https://www.europae-archaeologiae-consilium.org/making-the-case-for-development-led

European Archaeological Council (EAC), https://www.europae-archaeologiae-consilium.org

European Association of Archaeologists (EAA), https://www.e-a-a.org

European Commission (2016a): Communication on European Cloud Initiative - Building a competitive data and knowledge economy in Europe. COM/2016/0178 final, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52016DC0178

European Commission (2016b): Open Innovation, Open Science, Open to the World – a vision for Europe. European Commission. DG Research and Innovation. Luxembourg: Publications Office of the European Union, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=16022

European Commission (2018a): Commission Staff Working Document - Implementation Roadmap for the European Open Science Cloud. SWD(2018) 83 final, Brussels, 14.3.2018, https://ec.europa.eu/research/openscience/pdf/swd_2018_83_f1_staff_working_paper_en.pdf#view=fit&pagemode=none

European Commission (2018b): Second High Level Expert Group on the European Open Science Cloud: Prompting an EOSC in practice. Final Report and Recommendation, November 2018, Luxembourg: Publications Office of the European Union, http://dx.doi.org/10.2777/112658

European Commission: Digital Single Market policy: Open Science, http://ec.europa.eu/digital-agenda/en/open-science

Expert Group on FAIR Data (2018): Turning FAIR into reality. Final Report and Action Plan from the European Commission Expert Group on FAIR Data, November 2018, Luxembourg: Publications Office of the European Union,http://dx.doi.org/10.2777/1524

FAIRsFAIR Fostering FAIR Data Practices in Europe (H2020, 3/2019-2/2022),https://www.fairsfair.eu

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., Linek S. & Sauermann A. (2015): A Reputation Economy: Results from an Empirical Survey on Academic Data Sharing. DIW Berlin, Discussion Paper, No. 1454, February 2015, http://www.hiig.de/wp-content/uploads/2015/02/dp1454.pdf

Felicetti A., Williams D., Galluccio I., Tudhope D. & Niccolucci F. (2018): NLP Tools for Knowledge Extraction from Italian Archaeological Free Text. In: 3rd Digital Heritage International Congress, San Francisco, October 2018, doi: 10.1109/DigitalHeritage.2018.8810001; https://pure.southwales.ac.uk/files/3201918/DH2018_TEXTCROWD_Submission_Final2.pdf

Figshare (2017): The State of Open Data 2017, https://figshare.com/articles/_/5481187; Interactive visualisation: https://knowledge.figshare.com/articles/item/state-of-open-data-2017; Dataset: https://figshare.com/articles/State_of_Open_Data_survey_2017/5480710

Figshare (2018): The State of Open Data 2018, https://doi.org/10.6084/m9.figshare.7195058; Interactive visualisation, https://knowledge.figshare.com/articles/item/state-of-open-data-2018; Dataset: https://ndownloader.figshare.com/files/13320101;

Figshare (Digital Science), https://figshare.com

Frosini L., Bardi A., Manghi P. & Pagano P. (2018): An aggregation framework for digital humanities infrastructures: The Parthenos experience. In: CIRES*it*, 8(1): 33-50. http://doi.org/10.2423/i22394303v8n1p33

Garaba, Francis (2015): Dodos in the archives: rebranding the archival profession to meet the challenges of the twenty-first century within ESARBICA. In: Archives and Records, 36(2): 216-225, https://www.researchgate.net/profile/Francis_Garaba/publication/275328891

GEDE - Group of European Data Experts, https://www.rd-alliance.org/groups/gede-group-european-data-experts-rda

Geser, Guntram (2016): Towards a Web of Archaeological Linked Open Data. ARIADNE WP15 Study, October 2016, http://legacy.ariadne-infrastructure.eu/wp-content/uploads/2019/01/ARIADNE_archaeological_LOD_study_10-2016-1.pdf

Geser, Guntram (2019a): ARIADNEplus and community data repositories. Innovative solutions for sharing open archaeological data. In: CHNT'24 - International Conference on Cultural Heritage and New Technologies, Vienna, 2019, https://www.chnt.at/wp-content/uploads/ARIADNEplus-and-community-data-repositories.pdf

Geser, Guntram (2019b): Innovation and Impact of the ARIADNE Initiative, pp. 187-199, in: Richards J. & Niccolucci F. (eds.): The ARIADNE Impact. Budapest: Archaeolingua, http://doi.org/10.5281/zenodo.3476712

Getty Art & Architecture Thesaurus, https://www.getty.edu/research/tools/vocabularies/aat/about.html

Getty Vocabularies as Linked Open Data, http://www.getty.edu/research/tools/vocabularies/lod/index.html

Ghali, Walid (2015): The Stage of Manuscript Digitization Projects in Some Egyptian Libraries and Their Challenges, pp. 302-318, in: Click A.B. et al. (eds.): Library and Information Science in the Middle East and North Africa. Berlin: DeGruyter; available at: https://ecommons.aku.edu/uk_ismc_faculty_publications/1

Ghamouh N. & Boulahlib M. (2015): Cultural heritage digitization projects in Algeria: Case study of the National Library. In: IFLA Journal, 41(3): 272-278; https://www.researchgate.net/publication/282513458

GO FAIR (self-governed initiative and implementation networks), https://www.go-fair.org

GO-FAIR: FAIRification Process, https://www.go-fair.org/fair-principles/fairification-process/

Hasnain A. & Rebholz-Schuhmann D. (2018): Assessing FAIR Data Principles Against the 5-Star Open Data Principles, pp. 469-477, in: European Semantic Web Conference 2018: The Semantic Web, ESWC 2018 Satellite Events. Springer LNCS, 11155; preprint, http://hdl.handle.net/10379/14876

Horizon 2020 Work Programme 2018-2020. 4. European research infrastructures (including e-Infrastructures), INFRAIA-01-2018-2019: Integrating Activities for Advanced Communities, p.53, https://ec.europa.eu/research/participants/data/ref/h2020/wp/2018-2020/main/h2020-wp1820-infrastructures_en.pdf

HumaNum - Très Grande Infrastructure de Recherche des Humanités Numérique, http://www.huma-num.fr

ICCROM - International Center for the Study of the Preservation and Restoration of Cultural Property, https://www.iccrom.org

ICDI - Italian Computing and Data Infrastructure, https://www.icdi.it/en/

Internet Archaeology (journal), https://intarch.ac.uk

Kansa, Eric (2012): Openness and archaeology's information ecosystem. In: World Archaeology, 44(4): 498-520; preprint, http://escholarship.org/uc/item/4bt04063

KEA European Affairs (2019): Market Analysis of the Cultural and Creative Sectors in Europe, Brussels, https://keanet.eu/wp-content/uploads/ccs-market-analysis-europe-KEA-March2019.pdf

Kilmurry, Hayley (2019): Sudan Memory: The Conservation & Digitisation of the Sudanese Archives. In: Genusit.com, 20 June 2019, https://www.genusit.com/sudan-memory-conserving-promoting-sudanese-cultural-documentary-heritage/

Lake, Mark (2012): Open Archaeology. In: World Archaeology, 44(4: special issue), 471-478, http://dx.doi.org/10.1080/00438243.2012.748521

Libyan Antiquities at Risk (Society for Libyan Studies & University of Leicester), https://laar.le.ac.uk

Linked Conservation Data (UK A&HRC project), https://www.ligatus.org.uk/lcd/

Linked Open Data Cloud, https://lod-cloud.net

Mapping Memory Manager - 3M tool (FORTH-ICS), https://isl.ics.forth.gr/3M/

Mařík, Jan (2016): The organisation of Czech archaeology – a socialist legal system applied in a market economy, pp. 47-52, in: Florjanowicz, P. (ed.): When Valletta meets Faro. The reality of European archaeology in the 21st century. EAC Occasional Paper No. 11. Budapest: Archaeolingua; https://www.europae-archaeologiae-consilium.org/eac-occasional-papers

Markets&Markets (2019): Natural Language Processing Market, December 2019, https://www.marketsandmarkets.com/Market-Reports/natural-language-processing-nlp-825.html

Marlet O., Francart T., Markhoff B. & Rodier X. (2019): OpenArchaeo for Usable Semantic Interoperability, pp. 5-14. In: ODOCH 2019 - Open Data and Ontologies for Cultural Heritage workshop, Rome, Italy, 3 June 2019, http://ceur-ws.org/Vol-2375/paper1.pdf

Marwick, Ben et al. (2017): Open Science in Archaeology. In: The SAA Archaeological Record, September 2017, pp. 8-13; https://www.researchgate.net/publication/320068431

MASA - Mémoire des Archéologues et des Sites Archéologiques, http://masa.hypotheses.org

McKenna, Lucy (2018): Information professionals and linked data. In: InfoToday, 30 July 2018, https://www.infotoday.eu/Articles/Editorial/Featured-Articles/Information-professionals-and-linked-data-126598.aspx

McManamon F.P., Kintigh K.W., Ellison L.A., and Brin A. (2017): tDAR: A cultural heritage archive for twenty-first-century public outreach, research, and resource management. In: Advances in Archaeological Practice, 5, Special Issue 3, 238-249, https://doi.org/10.1017/aap.2017.18

MEGA Jordan (monuments database), http://www.megajordan.org

Mendeley Data (Elsevier), https://data.mendeley.com

Mitchell, Peter (2019): Carpe Diem: building African archaeology in the twenty-first century, pp. 7-15, in: di Lernia S. & Gallinaro M. (eds.): Archaeology in Africa. Potentials and perspectives on laboratory & fieldwork research. Arid Zone Archaeology, Monographs Series 8, Sapienza Università di Roma, Italy, https://www.academia.edu/40844865/

Moedas, Carlos (2015): Open Innovation, Open Science, Open to the World. Speech at the 'Opening up to an ERA of Innovation' Conference, Brussels, 22 June 2015, http://europa.eu/rapid/press-release_SPEECH-15-5243_en.htm

Mons B., Neylon C., Velterop J., Dumontier M., da Silva Santos L.O.B. & Wilkinson M. (2017): Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. In: Information Services & Use, 37(1): 49-56, http://content.iospress.com/articles/information-services-and-use/isu824

Mons, Barend (2016): We need 500.000 respected data stewards to operate the European Open Science Cloud: Interview with Barend Mons. In: e-IRG news blog, 4 May 2016, http://e-irg.eu/news-blog/-/blogs/we-need-500-000-respected-data-stewards-to-operate-the-european-open-science-cloud

Moraitou E., Aliprantis J., Christodoulou Y., Teneketzis A. & Caridakis G. (2019): Semantic Bridging of Cultural Heritage Disciplines and Tasks. In: Heritage 2019, 2(1): 611-630, https://doi.org/10.3390/heritage2010040

Munawar, Nour A. (2019): Cultural Heritage and the Arab Spring: A Review of (Inter)National Efforts to Safeguard Heritage Under Fire, pp. 83-115, in: Çakmak C. & Özçelik A.O. (eds., 2019): The World Community and the Arab Spring. Palgrave Macmillan, doi: 10.1007/978-3-319-60985-0

NAHAN - North African Heritage Archive Network, http://www.nahanweb.org

National Library of Israel, collections, https://web.nli.org.il/sites/nli/English/collections/

Near Eastern Archaeology, 2015, 78.3: Special Issue on The Cultural Heritage Crisis in the Middle East, https://www.journals.uchicago.edu/toc/nea/2015/78/3

Niccolucci F. & Richards J. (2019): ARIADNE and ARIADNEplus, pp. 7-25, in: The ARIADNE Impact. Budapest: Archaeolingua, http://doi.org/10.5281/zenodo.3476712

Nomisma (numismatics datasets and ontology), http://nomisma.org

OpenAIRE - Open Access Infrastructure for Research in Europe, http://www.openaire.eu

OSPP - Open Science Policy Platform (2018): OSPP Recommendations. European Commission, DG Research and Innovation. Brussels, April 2018, https://ec.europa.eu/research/openscience/pdf/integrated_advice_ospp_recommendations.pdf

Parcak, Sarah (2017): Moving from Space-Based to Ground-Based Solutions in Remote Sensing for Archaeological Heritage: A Case Study from Egypt. In: Remote Sensing, 9(12), 1297, https://doi.org/10.3390/rs9121297

PARTHENOS - Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies (H2020, 5/2015-10/2019), http://www.parthenos-project.eu

Pelagios Network, https://pelagios.org

PeriodO - Periods, Organized (University of Texas at Austin), http://perio.do

Perring, Dominic (2016): Working for commercial clients: the practice of development-led archaeology in the UK, pp. 95-104, in: Florjanowicz, P. (ed.): When Valletta meets Faro. The reality of European archaeology in the 21st century. EAC Occasional Paper No. 11. Budapest: Archaeolingua; https://www.europae-archaeologiae-consilium.org/eac-occasional-papers

Piron F., Diouf A.B., Dibounje Madiba M.S. et al. (2017): Le libre accès vu d'Afrique francophone subsaharienne. In: Revue française des sciences de l'information et de la communication [En ligne], 11|2017, http://dx.doi.org/10.4000/rfsic.3292

Pleiades (community-built gazetteer and graph of ancient places), http://pleiades.stoa.org

Pundit (digital annotation platform), http://thepund.it

Qatar Digital Library, https://www.qdl.qa/en

Randoin, Bernard (2016): French preventive archaeology: administrative organisation, role of the stakeholders and control procedures, pp. 59-64, in: Florjanowicz, P. (ed.): When Valletta meets Faro. The reality of European archaeology in the 21st century. EAC Occasional Paper No. 11. Budapest: Archaeolingua; https://www.europae-archaeologiae-consilium.org/eac-occasional-papers

Rayne L., Bradbury J., Mattingly D., Philip G., Bewley R. & Wilson A. (2017): From Above and on the Ground: Geospatial Methods for Recording Endangered Archaeology in the Middle East and North Africa. In: Geosciences (Switzerland), 7(4):100, October 2017, doi: 10.3390/geosciences7040100, https://www.researchgate.net/publication/320251994

Research Data Alliance: FAIR Data Maturity Model WG, https://www.rd-alliance.org/groups/fair-data-maturity-model-wg

Richards J. & Niccolucci F. (eds., 2019): The ARIADNE Impact. Budapest: Archaeolingua, http://doi.org/10.5281/zenodo.3476712

Riedel, Dagmar (2016): Of Making Many Copies There is No End: The Digitization of Manuscripts and Printed Books in Arabic Script, pp. 65-91, in: Muhanna, Elias (ed.): The Digital Humanities and Islamic & Middle East Studies. Berlin & Boston: De Gruyter; manuscript https://doi.org/10.7916/D82F7NB4

RISCAPE - Asmi A., Ryan L., Salmon E, Kubiak C., et al. (2019): International Research Infrastructure Landscape 2019. RISCAPE, December 2019, http://doi.org/10.5281/zenodo.3539254

RISCAPE - European Research Infrastructures in the International Landscape, https://riscape.eu

ROCEEH - The role of Culture in Early Expansions of Humans (Heidelberg Academy of Sciences and Humanities, Germany), http://www.roceeh.net

Schonfeld R. & Springer R. (2019): The Research Data Sharing Business Landscape. In: The Scholarly Kitchen, 2 October 2019, https://scholarlykitchen.sspnet.org/2019/10/02/research-data-sharing-business-landscape/

SEADDA - Saving European Archaeology from the Digital Dark Age (COST Action), https://www.seadda.eu

Smith-Yoshimura, Karen (2018): Analysis of 2018 International Linked Data Survey for Implementers. In: code4lib Journal, Issue 42, 8 November 2018, https://journal.code4lib.org/articles/13867

SSHOC - Social Sciences and Humanities Open Cloud (H2020 project, 1/2019-4/2022), https://sshopencloud.eu

tDAR - The Digital Archaeological Record (Digital Antiquity, Arizona State University, USA), http://www.tdar.org

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

Tenopir C., Dalton E.D., Allard S., Frame M., et al. (2015): Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide. In: PLoS ONE, 10(8): e0134826, 26 August 2015, https://doi.org/10.1371/journal.pone.0134826; data in S1 Appendix, https://doi.org/10.1371/journal.pone.0134826.s001

TII - Transport Infrastructure Ireland (2017): Digital heritage collections launch. TII website, 19 June 2017, https://www.tii.ie/news/archaeology/tii-digital-heritage-coll/

Vlachidis A. & Tudhope D. (2016): A knowledge-based approach to Information Extraction for semantic interoperability in the archaeology domain. In: Journal of the Association for Information Science and Technology, 67(5): 1138-52, doi: 10.1002/asi.23485; https://discovery.ucl.ac.uk/id/eprint/1556217/1/Vlachidis_A%20Knowledge-Based%20Approach%20to%20Information%20Extraction_.pdf

Vocabulary Matching Tool (Hypermedia Research Group, USW), http://heritagedata.org/vocabularyMatchingTool/

WGS84 - World Geodetic System 1984, http://earth-info.nga.mil/GandG/wgs84/

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

Wilson A.T. & Edwards B. (eds., 2015): Open Source Archaeology. Ethics and Practice. De Gruyter Open, http://www.degruyter.com/viewbooktoc/product/460080

Zaccaria, Massimo (2018): Digital Sources for the History of the Horn of Africa. In: Oxford Research Encyclopedias: African History, Mach 2018, http://dx.doi.org/10.1093/acrefore/9780190277734.013.341

Zerbini, Andrea (2018): Developing a Heritage Database for the Middle East and North Africa. In: Journal of Field Archaeology 43, Supplement 1, S9-S18, https://doi.org/10.1080/00934690.2018.1514722