D3.2 Interim report on policies and strategies

Version Final

Date 21th of December 2021

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 name, Title and Organisation:	**Prof. Franco Niccolucci, PIN Scrl - Polo Universitario "Città di Prato"**
Tel:	**+39 0574 602578**
E-mail:	**franco.niccolucci@pin.unifi.it**
Project website address:	**www.ariadne-infrastructure.eu**

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.

Author:	**Hella Hollander, DANS-KNAW**
Contributing partners:	**Holly Wright, UoY-ADS**
	Guntram Geser, SRFG
	Paola Ronzino, PIN
	Sheena Bassett, PIN
	Flavia Massara, MIBACT-ICCU
	Peter Doorn, DANS-KNAW
Quality control:	**Stephanie Williams, PIN**

Document History

- 4.11.2021 – Draft Version 0.1
- 7.12.2021 – Draft Version 0.2
- 21.12.2021 – Final Version

Table of Contents

1 Executive Summary

ARIADNEplus objectives

As a follow up on the initial ARIADNE project, the overall objective of ARIADNEplus is to serve archaeological researchers and data management communities by proceeding to improve data sharing and the (re)use of data resources, which are dispersed through Europe and often difficult to discover and access as different silos (institutional, national or disciplinary) still exist.

Policies and Good Practices for FAIR Archaeological Data Management

This interim report, "Policies and Good Practices for FAIR Data Management", is the second deliverable of Work Package 3. It describes the activities carried out by the different partners during the first 2 and a half years of the ARIADNEplus project, as well as the results achieved through the work package. The following partners are involved: DANS-KNAW, PIN, UoY-ADS, CNR, CONICET, BUP, NIAM-BAS, AMZ, ARUP, AU, UH, CNRS, INRAP, RGK, ATHENA-RC, PP, HNM, FI, IAA, MIBACT-ICCU, NARA, DGPC, SND and ASU. The first and initial report D3.1 takes into account what was achieved in projects that preceded ARIADNEplus. The outcomes of the initial ARIADNE project as well as the PARTHENOS project were used as a starting point to build on.

The objectives of Work Package 3 Policies and Good Practices for FAIR Data Management are to:

- Support the creation of FAIR data in the archaeological sector
- Define and spread guidelines to good practices in archaeological data management
- Adapt standard quality criteria for datasets and data to the archaeological case, and support their implementation among users.

Chapter 2 describes how to define and spread guidelines to good practices in archaeological data management. Commonly developed and widely applicable guides will ensure that archaeological data will be FAIR and available in the long-term.

Chapter 3 presents activities to develop and implement a portfolio of tools to support users in their work with archaeological data. The activities described in this period focused on the integration and alignment of the ARIADNEplus DMP Researcher Template for Archaeological Datasets with a Domain Data Protocol based on the Science Europe core requirements for Research Data Management.

Chapter 4 shows the importance of sharing experiences from partners with already certified repositories to partners willing to set up an archaeological data repository. Providing guidelines and support on repository creation and management is the focus of activity here.

Chapter 5 describes what partners willing to certify their repository need to be provided with: the explanation of and training on accreditation requirements when applied to repositories of archaeological data with a perspective on international initiatives. Achieving a Trustworthy Data Repository status, and making and keeping data FAIR is a joint journey.

Chapter 6 highlights the application of the FAIR principles to archaeological data, taking into account different regulations throughout Europe and the potential sensitivities and IPR-related issues. A survey was held to collect and analyse information for assessing the current policies that determine access to and reuse of data held by digital archaeological repositories, and providing guidance and support needed to make the repositories and data FAIR.

Chapter 7 describes training activities on FAIR Data Management.

2 Good practices in archaeological data management

Task 3.1 builds on the work of the first phase of the ARIADNE project under WP4 *Good Practices and Dissemination,* and specifically on Task 4.5 *Good Practices* and Task 4.6 *Guides to Good Practice*. UoY-ADS leads the task, with SND, MIBACT-ICCU, and DGPC.

The initial survey of ARIADNE partner organisations carried out as Task 4.5 in the first phase of the ARIADNE project highlighted the existence of a variety of guidance and Good Practice documents. These documents reflected a broad range of expertise and function while also highlighting several specific themes which formed the objectives for work carried out under Task 4.6 Guides to Good Practice. The objectives included:

- The alignment and referencing of existing Good Practice documents.
- The creation of case studies illustrating the application of Good Practice documents to specific data sets for which no good practice currently exists.
- The incorporation of guidelines produced by the ArchaeoLandscapes and 3D-ICONS projects into ARIADNE guidelines, and the illustration of these guidelines through relevant case studies.
- The revision, creation, or enhancement of guidelines for 3D datasets.
- The creation of guidelines for data from scientific dating and analysis, specifically dendrochronological datasets.

Tasks 4.5 and 4.6 successfully met these objectives and produced several new and much-needed guidelines which individually incorporate one of more of the areas identified for contribution. The guides and case studies successfully incorporated existing material and guidelines from a wide range of sources, ranging from the outputs of other collaborative projects such as 3D-ICONS through to organisation-specific guidelines produced by project partners such as DAI and DANS. Additionally, case studies were used both within individual guides and as stand-alone contributions to successfully illustrate the application of data selection, archiving, and documentation procedures to real-world datasets. When viewed together, the outputs of Task 4.5 and 4.6 highlighted that, while language, procedure, and archaeological practices may vary widely between countries and institutions, the data that arises from archaeological investigations and projects, irrespective of geography, share common elements that allow guides for good practice to be commonly developed and widely applicable.

ARIADNE/ARIADNEplus representation and results in other projects

Overlapping with the end of the first ARIADNE project and the start of ARIADNEplus, the Preparatory Phase for the European Infrastructure for Heritage Science (E-RIHS) was underway and management of archaeological science data was in charge of the partner ADS within such project. The work done for E-RIHS resulted in a report that reviewed issues concerning good practice data management for heritage science[1]. The intention was to provide a policy framework to be implemented by E-RIHS, but the report was designed to better understand data management workflows within the heritage science domain. It was found to be a broad and heterogenous area, and the report was only able to include a representative range of approaches such as materials analysis, dating methods, archaeological science, biomolecular archaeology, synchrotron methods, and conservation science, amongst others.

[1] https://zenodo.org/record/3946202

The report followed the framework provided by the FAIR principles (Findable, Accessible, Interoperable and Re-usable) but interpreted them in the context of heritage science. Examples were given with reference to a substantial appendix which covered a broad range of heritage science data types. E-RIHS was also involved in the development of the DARIAH Heritage Data Reuse Charter, to complement the FAIR principles, and create both principles and mechanisms to which heritage science practitioners should adhere. It was recommended that E-RIHS researchers complete a Data Management Plan as a condition of support for their usage of an E-RIHS facility, and this would assist E-RIHS researchers, facilities and repositories in following FAIR data recommendations.

With the completion of the E-RIHS Preparatory Phase project, ADS undertook discussions as to whether to pursue another Guide to Good Practice for Archaeological Science Data. Based on the lessons learned from participation in E-RIHS, this was deemed impractical at the current time. This area of archaeological research was deemed so complex and diverse, that any single Guide would not be of practical use. Instead, work continues within ARIADNEplus on trying to better understand the archaeological science data landscape, so that it may be possible to create guidance in the future.

Also overlapping with the start of ARIADNEplus is the Social Science and Humanities Open Cloud (SSHOC) project. SSHOC is funded by the EU framework programme Horizon 2020 and includes 20 partner organisations and their 27 associates to develop the social sciences and humanities area of the European Open Science Cloud (EOSC), including ADS from ARIADNE. This project has been undertaking an examination of the issues and challenges faced in providing FAIR access to archaeological data as resulting from the ARIADNE research.

Much was learned about the complexity of undertaking an audit on the ADS archiving policy at the aggregation level, and it was deemed a worthy undertaking to provide FAIR feedback to ARIADNE metadata providers, but it also required additional capacity from ARIADNE partners which was not available at the time and is not planned in the current project description of work. It is hoped this extra work may still be possible before the end of the current phase of ARIADNEplus, but this additional activity must be weighed alongside other core project commitments. As such, this capacity will be reassessed in the first half of 2022.

Harmonized Guidance for Protocol and Data Management Plan templates

Science Europe, the European association representing the interests of major public research performing and research funding organisations, published two documents with great relevance for Task 3.1 on good practices in archaeological data management. The Guidance Document “Presenting a Framework for Discipline-specific Research Data Management” of 2018 proposed the creation of domain-specific protocols to be used as standardised templates for RDM, reducing the administrative burden on both researchers and research organisations, as well as on funders[2].

In 2019 the “Practical Guide to the International Alignment of Research Data Management” was prepared by experts from Science Europe Member Organisations[3]. This guide aims to align RDM core requirements across research and funding organisations. Many research councils and universities in Europe, including the Horizon Europe programme, accept the Science Europe core requirements as the basis for a data management plan (DMP) in order to make research data open and FAIR.

[2] http://www.scienceeurope.org/our-resources/implementing-research-data-management-policies-across-europe/

[3] http://www.scienceeurope.org/our-resources/practical-guide-to-the-international-alignment-of-research-data-management/

It is also gaining ground beyond Europe, as is witnessed by a recent translation into Chinese[4]. The Science Europe guidance received an update in 2021, when an Evaluation Rubric[5] was added, providing further guidance for evaluators of DMPs on what are considered to be necessary or acceptable answers[6]. A close collaboration between Science Europe and ARIADNE was agreed, with DANS and PIN (both ARIADNE partners and the former also a key partner of Science Europe) in charge of the work.

Work had already started in 2019 to develop protocols for various domains on the basis of these documents. A first draft of a Domain Data Protocol for Social and Behavioural Sciences was presented at the European Survey Research Association (ESRA) Conference (Zagreb, July 2019)[7], with accompanying guidance oriented to this domain. In the ARIADNEplus project, a Domain Data Protocol for Archaeology, compliant with the Core DMP requirements by Science Europe, was developed as a complement to the already existing DMP template for archaeological research data management (see section 3).

In the first 18 months of the ARIADNEplus project, the activities focussed on the definition of the guidance, including the PARTHENOS Guidelines on how to make data FAIR[8], the guidance drafted in collaboration with the OpenAIRE project and cross-referenced with external resources, like the ones available under the "training hub" section of the ARIADNEplus website. This enables to create links that support researchers and institutions in developing a DMP adopting standards and procedures that are shared among the different groups representing the different disciplines of the archaeological domain (See Annex 1 of D3.1).

The work carried out by the working group formed by Peter Doorn (DANS) and Paola Ronzino (PIN) in the second reporting period (M19-36) continued with the refinement of the ARIADNEplus DMP template's guidance and on the harmonization with the core requirements formulated by Science Europe, incorporating the domain protocol-idea by proposing norms for good practices in data management that can be generally accepted by the archaeological community. The Protocol was also checked against the Science Europe Evaluation Rubric for DMPs. The Protocol provides extensive guidance, with explanations of and references to relevant information sources concerning archaeological data management.

[4] https://data.depositar.io/en/dataset/se_rdm_guides
[5] https://zenodo.org/record/4915862#.YcBAqhPMJsM
[6] https://doi.org/10.5281/zenodo.4915862
[7] https://bit.ly/31mQAbp
[8] https://www.PARTHENOS-project.eu/portal/policies_guidelines

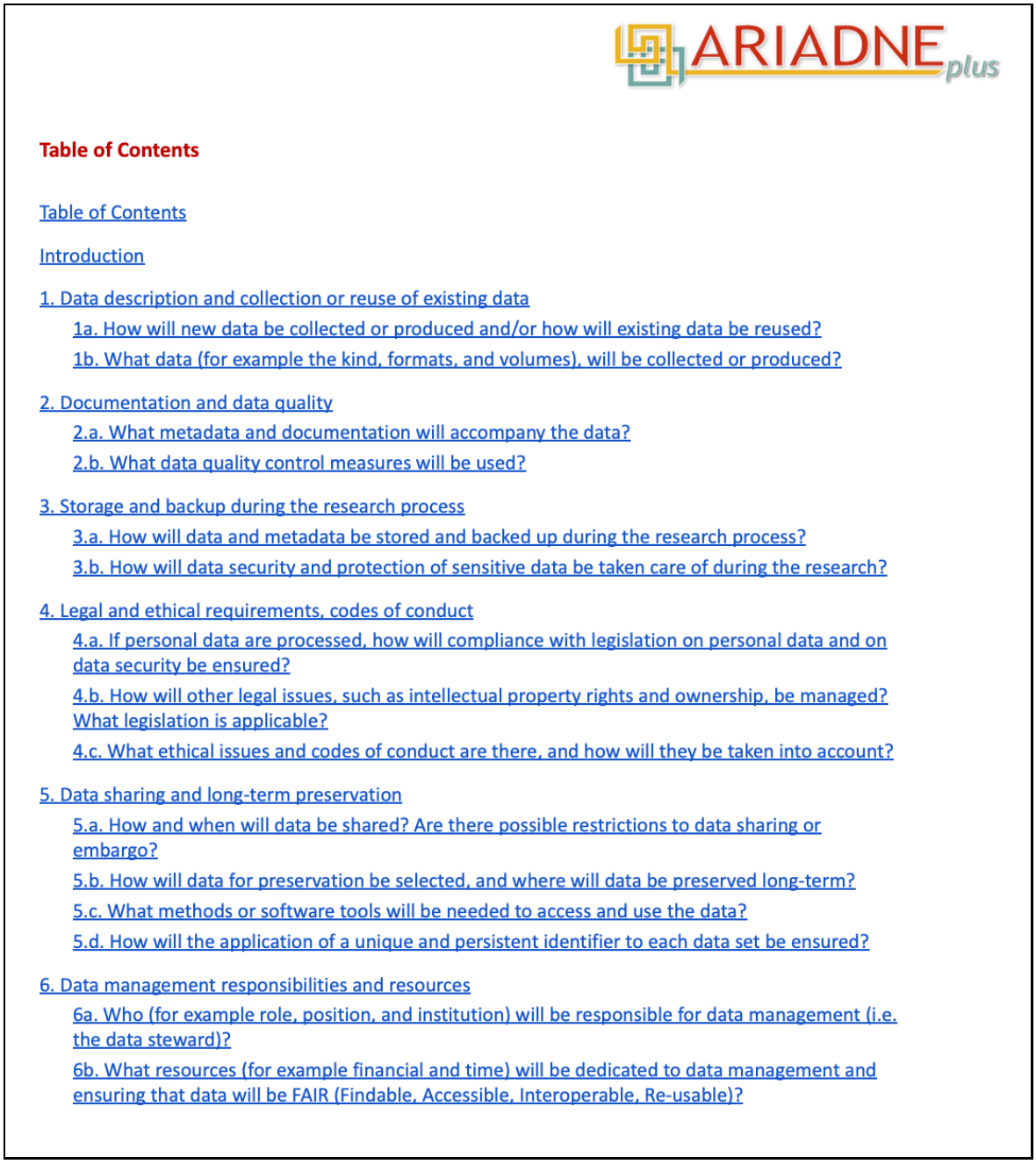

Table of Contents

Figure 1 List of Contents of the Data management Guidance document (conforming SE Core Requirements)

The guidance document follows the order of the Science Europe core requirements, and it is possible to link back and forth between the templates, protocol and guidance. This guidance document obviously makes use of previously existing guidelines and good practices, but it is exactly the tailoring to the individual items of the protocol and DMP templates that sets it apart from other guides. The full guidance document of about 20 pages is available in the training hub of the ARIADNEplus website[9].

[9] https://training.ariadne-infrastructure.eu/dmp-guidance/

3 ARIADNEplus policy support tools

Task 3.2 will implement a portfolio of tools created to support users in their work with archaeological data. DANS-KNAW leads the task, supported by PIN, MIBACT-ICCU, DGPC and other partners as required.

One of the objectives of this task is the release of a tool that enables the archaeological community to comply with the requirements of funding institutions, who often request to submit a data management plan (DMP) to document their research process, demonstrating that they manage their data responsibly.

As part of the task, ARIADNEplus has developed a DMP template for the archaeological community providing full indications on the compilation of all the questions addressed, which are accompanied by a set of guiding statements, and relevant responses provided to assist users in completing their DMP with an online tool. The ARIADNEplus template builds on the work done in the PARTHENOS project[10] and is compliant with the H2020 DMP template. PIN is the partner responsible for the development of such a template and of the online tool, with the support of DANS-KNAW.

The collaboration between ARIADNEplus and OpenAIRE has been established through joint activities carried out by the team responsible for the development of the ARIADNE DMP tool and the OpenAIRE group involved in the development of the ARGOS tool. The content of the DMP templates, both compliant with the H2020, has been mapped to each other to identify possible gaps and to join forces in the implementation of guidelines useful for filling the templates with the right information, besides referencing the resources already published by both projects.

The work on the refinement of the ARIADNEplus DMP template's guidance and on the harmonization with the core requirements formulated by Science Europe is described below. Further work included the alignment of the ARIADNEplus template with the new specifications required by Horizon Europe to identify possible gaps and the implementation of a tool that will incorporate the three templates.

The ARIADNEplus DMP template for Archaeological Datasets

Based on the Open Science initiative and the FAIR principles, the ARIADNEplus DMP template for Archaeological Datasets addresses researchers in the archaeological domain and is tailored to the community needs, including standards and tools commonly used in their daily practices. It satisfies the needs of research organizations that manage institutional repositories, with a section specifically tailored for them, and of researchers, as they are both data producers and data users, each having a specific perspective on data quality and FAIRness issues. Furthermore, the ARIADNEplus DMP is aimed at researchers and institutions that still don't have any written policies on how to write a clear document that explains what data they will create, how it will be managed and what their plans are for data sharing and preservation.

With its structure and the suggested answers, the DMP helps researchers think about what to do with their research data, how to collect and to keep track of it, thus helping to identify the support, standards and services needed. It is a useful instrument to plan for short- and long-term storage, and to prepare data for re-use by acknowledging the sources and intellectual contributions according to legal terms and conditions that may include limited privileged use.

[10] https://www.parthenos-project.eu

The ARIADNE DMP template is compliant with the guidelines on FAIR Data Management published by the European Commission to ensure that research data is publicly available, to help Horizon 2020 beneficiaries in making their research data findable, accessible, interoperable and reusable, with the main objective of increasing the scientific quality of the funded projects and to foster their replicability. The ARIADNE DMP template follows the same structure of the Horizon 2020 Guidelines (deemed mandatory in the template), and includes a set of questions organized into the following sections:

1. Data Summary
2. FAIR Data
3. Allocation of resources
4. Data Security
5. Ethical aspects
6. Other

By answering the questions included therein, researchers will provide, among other things, information on:

- Data set description: with detailed information on the scientific focus and technical approach.
- Standards and metadata: users can select among several protocols and standards adopted by the archaeological community or describe the proprietary schema used to structure their data so that other scientists can make an assessment and reproduce the dataset.
- Name and persistent identifier for the data sets: guarantees using repositories that will provide a unique and persistent identification (an identifier) for their data sets and a stable resolvable link where their datasets can be directly accessed.
- Curation and preservation methodology: providing information on the standards that will be used to ensure the integrity of their data sets and the period during which they will be maintained, as well as how they will be preserved and kept accessible in the longer term. If available, users can provide a reference to the public data depository in which their data will be deposited.
- Data sharing methodology: providing information on how their data sets can be accessed, including the type of license under which they can be accessed and re-used, and information on any restrictions that may apply. Users can specify and justify the timing of data sharing, for example, as soon as possible after the data collection, or at the end of the project. In the section on making data reusable, researchers can find out more about Open access policies that promote research data sharing and practical suggestions for managing IPR issues.

All the sections above include questions aimed at the researchers, while section 4 on “Data Security” mostly addresses data managers and repository managers as it concerns information on data recovery, as well as secure storage and transfer of sensitive data, details that a researcher is not necessarily informed about.

The need for support in the compilation of a DMP was strongly expressed by a group of experts that responded to a survey carried out by the Research Data Management team of the OpenAIRE project and the FAIR Data Expert Group[11] to collect feedback for the evaluation of the Horizon2020 approach

[11] M. Grootveld, E. Leenarts, S. Jones, E. Hermans, and E. Fankhauser, OpenAIRE and FAIR Data Expert Group survey about Horizon 2020 template for Data Management Plans (Version 1.0.0), 2018 [Data set]. Zenodo. http://doi.org/10.5281/zenodo.1120245

to DMPs in order to identify gaps and collect suggestions for improvement. This need has been further confirmed by the extended community of ARIADNEplus and by the archaeologists and experts of the archaeological domain that are part of the SEADDA community, to which we submitted the DMP template and the tool, asking for their comments and validation.

Domain Protocol for Archaeological Data Management

As part of task 3.2, a Domain Data Protocol for Archaeology, compliant with the Core DMP requirements by Science Europe, was developed as a complement to the already existing DMP template for archaeological research data management. The current version of the domain protocol is available here: https://www.surveymonkey.com/r/ZSCL7NV.

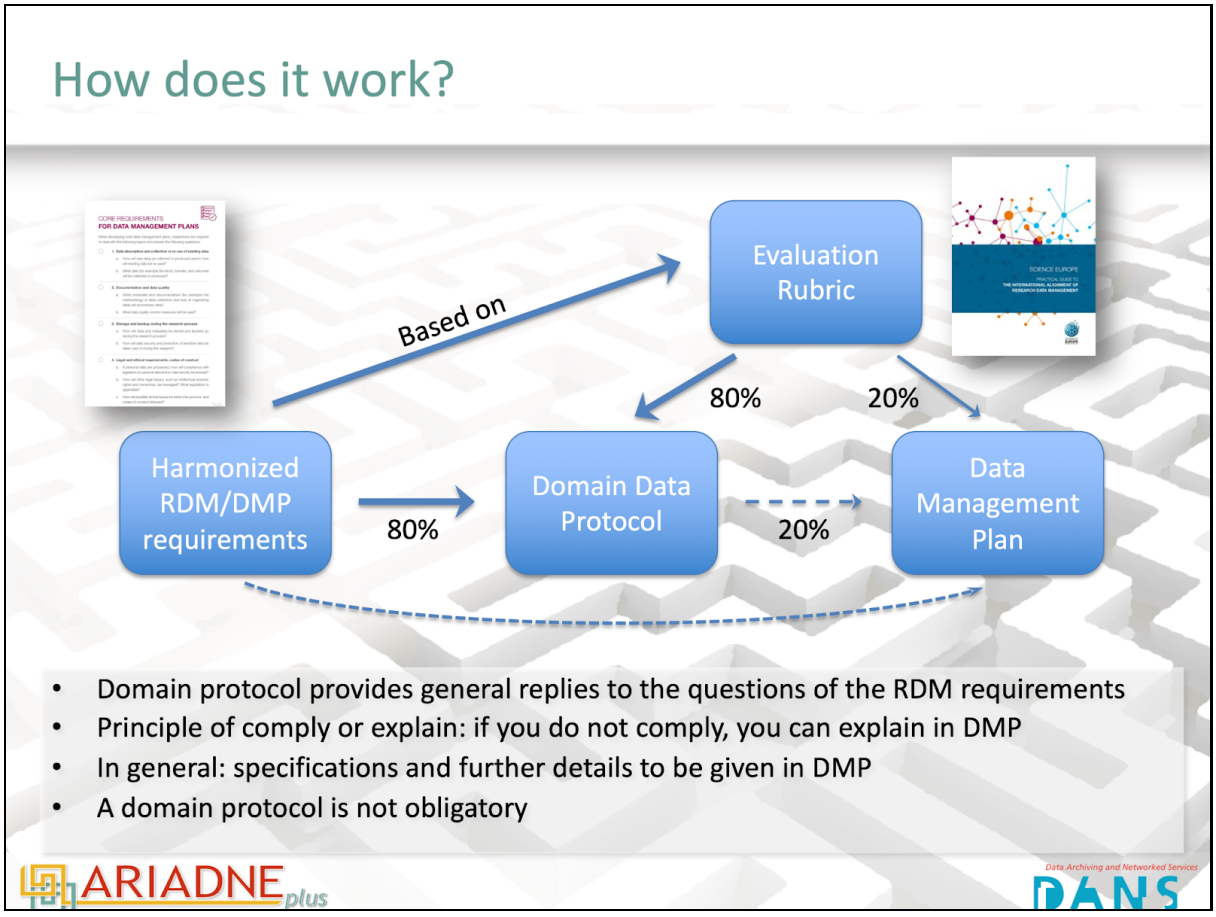

Figure 2 How a domain protocol makes data management planning more efficient

This Protocol offers pre-formulated statements and replies to the topics raised and questions asked in the core requirements for research data management (RDM) as formulated by Science Europe, and also incorporates the suggestions of the Evaluation Rubric. The Protocol is obviously tailored to the field of archaeology and heritage studies, and is based on the principle of "comply or explain". Archaeologists complying with the statements in the Protocol can save a lot of time when preparing a DMP: a motivation is required only when deviating from a standard reply. The protocol will link to the relevant question in the Horizon Europe template for archaeology. It is also possible to provide extra information on the standard answers of the protocol in free text boxes, which can be filled out with optional further explanations.

Figure 3 Screenshot of the first item of the Archaeology data protocol. In case of non-compliance, the protocol will link to Horizon Europe template for archaeological datasets

Diverging RDM requirements of Science Europe and Horizon Europe

As the call for harmonization of DMP requirements across research funding organisations came from the Director-General of DG Research and Innovation (RTD) at that time, and as representatives of RTD supported the Science Europe core requirements, we had hoped and expected that the DMP requirements of the Horizon Europe programme would closely follow the Science Europe core requirements. However, when the first draft of the HE requirements became available in May 2021, it appeared that these were largely a continuation of the Horizon 2020 requirements, but with a number of modifications.

Although the aims and topics of both the Science Europe and the Horizon Europe requirements are quite similar, they differ substantially in detail. The structure and order of the topics and questions to be covered, as well as many formulations, differ. The emphasis on topics and the way in which these are grouped also differ. These discrepancies are unfortunate, as the continuation of the differences is confusing for researchers and hinders the creation of uniform DMP templates that can be used for multiple purposes. We have discussed the situation directly with representatives of both Science Europe and of the Open Science Unit of RTD, but for the time being the maximum attainable seems to be that the format of DMPs for Horizon Europe projects is not prescribed, and thus the order of subjects according to Science Europe (and hence the protocol) is also acceptable.

Mapping between H2020 and Horizon Europe templates

To solve the situation for the DMP tools and guidance for archaeologists in the ARIADNEplus project in the best possible way, we made a mapping of the three sets of requirements (H2020, Horizon Europe and Science Europe) to provide clarity on both similarities and dissimilarities[12]. We decided to offer archaeologists a choice of approaches and online tools to assist them in making a data management plan. We will now be offering the following:

1. The already existing **ARIADNEplus DMP Researcher Template for Archaeological Datasets**[13] based on the Horizon 2020 requirements, initially developed in the PARTHENOS project[14].
2. The new **Horizon Europe Template for Archaeological Datasets.**
3. A **Protocol for Archaeological Data Management**, based on the Science Europe guide for research data management (including directions for evaluating DMPs). The protocol will link to the HE template for archaeological datasets when deviating from the provided standard answer (non-compliance) or when additional explanations are needed.
4. A harmonized **Guide for Archaeological Data Management Planning** that can be consulted for both DMP templates and the protocol[15].

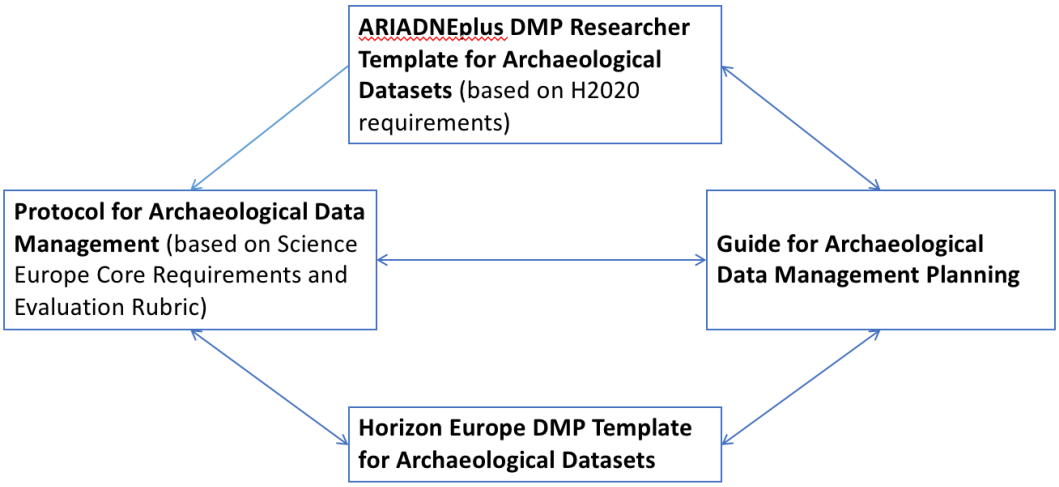

Figure 4 Links between the ARIADNEplus data management tools and guidance

The ARIADNEplus DMP tool

To facilitate the compilation of the ARIADNEplus DMP Researcher Template for Archaeological Datasets, an ad hoc application has been developed by PIN, the design of which has taken into consideration both the practical needs of researchers, and the current technological evolution that digital documents are undergoing. Stored on PIN's server, the tool is accessible from the Services page of the ARIADNE portal[16].

As described above, the need for many researchers to compile a DMP to be submitted to the funding agencies to verify compliance with the requirements set by Science Europe, and the publication by

[12] Mapping between H2020-HE-SE: https://tinyurl.com/2p8j77r4
[13] https://vast-lab.org/dmp/ariadneplus/form/index.html
[14] https://www.parthenos-project.eu/portal/dmp
[15] https://training.ariadne-infrastructure.eu/dmp-guidance/
[16] https://portal.ariadne-infrastructure.eu/services

the European Commission of a new DMP template for projects funded by the Horizon Europe programme for 2021-2027 prompted us to review the design of the tool already implemented for the ARIADNEplus DMP.

The tool now consists of three modules that can be selected from the home page, based on the user's needs (Figure 5). Once the relevant template is selected, the user is directed to the pertinent page and guided to fill in the questions by following the links to the guidance document. The online interface has been designed to facilitate the compilation of the DMP through the use of intuitive and user-friendly solutions. The questions that the researcher is invited to answer are divided into successive pages, enriched by a common progress bar that presents itself as the main reference point for the user. The overall view of the various parts that make up the model guides the user step by step, indicating approximately the time required to conclude them.

Figure 5 Home page of the ARIADNEplus DMP tools

Each page groups similar thematic questions, divided into mandatory and optional, enriched by informative pop-ups that help the user fill them in. If some of the points deemed mandatory for submitting the DMP have not been completed, their number will be displayed in red in the progress bar. At the end of the compilation procedure it will be possible to download the information it contains in PDF format, in TXT and in JSON. The JSON file is essential within the application, as it offers the user the opportunity to save a version of the work. In fact, the compilation of the questionnaire can be interrupted at any time by downloading the JSON file containing the current data. This file can be reloaded within the online interface to continue and finish the job. If instead the compilation of the questionnaire is definitive, the data contained in the file will constitute a version of the DMP useful for any subsequent revisions or updates. None of the user's personal data are collected nor stored.

Each template can be completed independently and saved as a single document. Nevertheless, since the level of detail of each template is different, for example from a more general view of the Domain Protocol to the more detailed analysis of the ARIADNEplus DMP, interlinks are created between the templates to learn more on a specific topic (see Figure 4).

Our final goal is to obtain a machine-actionable DMP, whose information can be automatically processed and understood by computers, and which is at the same time interoperable, editable and shareable within the community of stakeholders. The design of future application developments is aimed at making the data contained within the DMPs shareable and interoperable between those research communities that will adopt common solutions to facilitate cooperation between their systems. Making documents interoperable means making sure that the information they contain can be exchanged between different systems in a complete and reliable way. For this it is necessary to

consider both the syntactic and the semantic aspects of the data. Computers can process and manage most of the information syntactically, if it is encoded in standard formats such as XML or JSON, but they are unable to interpret and "understand" it if it is not modelled using controlled vocabularies and standards. The DMPs that can be generated with the current version of the tool already meet the requirements for syntactic interoperability, thanks to the encoding in JSON format.

The machine actionable version of the ARIADNEplus DMP Researcher Template for Archaeological Datasets template is currently under definition as it relies on the mapping to the AO-CAT model, the CIDOC CRM extension developed for the interoperability of the datasets integrated by ARIADNEplus. This will offer researchers the opportunity to benefit from sharing information.

The group responsible for this activity collaborated with a team from the OpenAIRE project, and integrated the ARIADNEplus DMP template into the ARGOS tool[17]. ARGOS is an open extensible service that simplifies the management, validation, monitoring and maintenance of DMPs. It allows researchers, managers, supervisors, etc. to create actionable DMPs that may be freely exchanged among infrastructures for carrying out specific aspects of the data management process in accordance with the intentions and commitment of data owners. The ARIADNEplus DMP template has been tested and embedded into the ARGOS environment. This guarantees greater visibility and to reach a larger community. A mapping of the ARIADNEplus DMP template with the RDA DMP Common Standard for machine-actionable Data Management Plans has been carried out to allow integration into the ARGOS environment. On the other side, ARIADNEplus contributed to the ARGOS multidisciplinary aspect by sharing standards, vocabularies and other information specific to the archaeological domain.

The added value of the ARIADNE DMP template compared to other existing templates, stands in the guidelines provided in support of the questions and the suggested answers based on the standards and operative workflows adopted in archaeology. This way, users have a better understanding of the processes and methodologies used, and may also consider possible alternatives to their research approach.

The Policy Wizard Tool

D3.1 describes the Policy Wizard Tool[18] as an online service to help archaeologists discover data policies and best practices. In this reporting period DANS and CNR started to revive the tool by updating the database behind it and by updating the code. The next step will be to add missing policies to it by approaching the ARIADNEplus network.

To guarantee a coherent dissemination of these valuable tools, webinars and training workshops will be organized with the aim to raise awareness on open research in archaeology and in the digital humanities sector. Furthermore, the outlines of the DMP templates and the guidelines will be made available to project partners so that translation into different languages may be produced to provide national versions to those countries that have not yet implemented their own template, e.g. Italy, Spain, Greece, etc.

[17] https://www.openaire.eu/argos/
[18] https://www.parthenos-project.eu/portal/wizard

4 Providing guidelines and support on repository creation and management

Task 3.3 provides guidelines and supports partners willing to set up an archaeological data repository. UoY-ADS leads the task, with SND and CNR-ISTI (NEMIS-Infra). Other partners will advise on national/local opportunities.

The COST Action Saving European Archaeology from the Digital Dark Age (SEADDA), continues to run concurrently with ARIADNEplus, allowing an increased focus on data stewardship and expansion of international participation for ARIADNEplus partners, and beyond. SEADDA[19] has participants from 34 countries, including all countries represented by ARIADNEplus partners. ARIADNEplus and the ARIADNE Portal continues to develop the state-of-the-art for the aggregation of archaeological data, and SEADDA focusses on the long-term trajectory of the data itself. It is meant to expand the capacity-building necessary for organisations, regions and countries to expand their participation in ARIADNEplus in a more equitable way through collaborative stewardship.

Task 3.3 was updated from draft guidance to be produced towards the middle of the project (with a revised version planned for delivery at the end) linked to TNA from WP9, to take advantage of SEADDA being funded concurrently with ARIADNEplus, and continues to allow deeper work and collaboration across four different working groups (WG), and the ability of ARIADNEplus partners to participate in SEADDA Short Term Scientific Missions for more intensive work. Three of the SEADDA Working Groups are specifically tasked with providing support for repository creation and management, the composition of which are a direct response to the lessons learned in the first phase of ARIADNE. Over the last 12 months, 28 articles were published in 2021 in a comprehensive, open access themed issue *Digital Archiving in Archaeology: The State of the Art*[20] representing a wide range of countries, nations and regions on the current state of data stewardship in archaeology.

A workshop to discuss the survey findings is planned for May 2022, in Braga, Portugal, which will result in a synthetic publication in 2023. SEADDA is also supporting stewardship through workshops held in national languages in a range of countries. These are now planned to take place in 2022 in Serbia, Ireland, France, Portugal and Turkey. Participation in a sustainable resource for supporting repository management called Community Owned digital Preservation Tool Registry (COPTR)[21] is also underway. COPTR is both a registry of preservation tools and a documentation platform for preservation workflows. In the coming year, members will participate in two COPTR hackathons focussed on 3D data and spatial data, and are considering a preservation workflow hackathon for established repositories of archaeological data. DANS has also conducted interviews with Research Infrastructures and other international stakeholders (e.g. DARIAH, E-RIHS) to undertake future watch outside the archaeology domain. The results will be used to present the state of the art of archaeological archiving and dissemination in Europe. In the coming year, an overview will be given of recent best practices and goals, including FAIR RDM, certification, preferred formats and an archaeological DMP and domain protocol and a workshop organised in collaboration with WG2 on tools, knowledge transfer and communication.

[19] http://seadda.eu
[20] https://doi.org/10.11141/ia.58.23
[21] https://coptr.digipres.org/index.php/Main_Page

ARIADNE partners from ADS and DANS also assisted in Project Urdar[22], headed by Uppsala University and the Swedish National Heritage Board (NHB), and financed by Riksbankens Jubileumsfond. Project Urdar is aimed at developing methods to export the data in circa 3600 projects created by the archaeological unit at the NHB (RAÄ UV) between 2000-2014 using the proprietary archaeological project management software called Intrasis. Sweden is currently lacking clear processes for the delivery and archiving of archaeological documentation data, and there is only a system in place for delivering reports to a national archive (Swedish National Heritage Board-SNHB). There is increased pressure from contract archaeology and County Administrative Boards to create a system for digital documentation, and to make sure valuable data is not lost over time, as it is now kept on servers at the individual companies with no guarantee of future sustainability.

Data created using Intrasis is not preservable in the long term. Through Urdar, data is to be preserved as close to the original as possible in open formats, but also be made findable and accessible through the interfaces of the NHB e-archive (Arkivsök) and the Historic Environment Record (Fornsök). As the project is now at the midpoint, two webinars were organised with the aim of presenting the progress and insights made, as well as to take the opportunity to increase the knowledge about good digital practice and FAIR data principles among the major producers and users of archaeological documentation. The seminars were attended by between 80-100 participants from contract archaeology, regional and national government agencies, museums and research institutes. The first seminar presented the project, explained FAIR data and highlighted the potential of international reusability through ARIADNEplus. The second seminar offered a look at the broader world and included presentations about the systems in place in the UK, with a presentation by ADS and the Netherlands, with a presentation by DANS. The history and current status of archiving stewardship practices at the RAÄ UV and SNHB can be found in the SEADDA publication *A Turn Towards the Digital. An Overview of Swedish Heritage Information Management Today*[23].

This autumn, the Dutch Cultural Heritage Agency (RCE) and DANS have launched DataverseNL as the new DCCD portal for depositing, downloading and (inter)nationally disseminating tree-ring data. DCCD stands for Digital Collaboratory for Cultural Dendrochronology. The DCCD is an initiative of the RCE and DANS and was partly funded by the Netherlands Organisation for Scientific Research (NWO). This international digital data library of tree-ring data contains measurements of tree-ring patterns (growth ring series) as well as their descriptive and interpretive metadata and research reports. By means of dendrochronological research the calendar year in which a tree was previously cut down can be determined. This allows wooden objects to be dated. In addition, more can be learned about development and use of landscapes. As a result, research reports and measurement series can be found about objects and sites from prehistory up to now. The archive contains data from European institutes. Each organisation displays its own collection of data related to archaeological sites, shipwrecks, buildings, furniture, paintings, sculptures and, for example, musical instruments.

As the DCCD portal was getting outdated and obsolete it was no longer a sustainable application. The RCE and DANS made the important decision last year to renew the DCCD portal and phase out the old one. In order to make the data of the current DCCD findable, accessible, and to be able to add new data, it was decided to continue the functionality of the DCCD in DataverseNL. This is a shared service provided by participating institutions. DataverseNL uses Dataverse software developed by Harvard University, which is used worldwide. To be able to migrate all data it was no longer possible for organisations to deposit their data in the old portal, but downloading data was still possible.

[22] https://www.arkeologi.uu.se/Research/Projects/urdar-en/
[23] http://dx.doi.org/10.11141/ia.58.19

Depositors were contacted about this and they were asked to agree to update their old user-licence to a formal and open access CC-by license. The depositors opted mostly for CC BY-NC-ND 4.0: https://creativecommons.org/licenses/by-nc-nd/4.0/ Furthermore, Dataverse depositors are able to determine who gets access as users can be given different rights and roles. A mapping of metadata fields from the old metadata schema to the new schema was investigated, discussed and executed. In the new system, the data is given a persistent identifier (DOI) making it easy for others to find and refer to the dataset by the citation information provided. A new feature is that updated datasets can be stored during research, making it possible to keep track of changes with version control. Not all functionality returned: as an old application and being a database using non-preferred formats (MS Access), the functionality of Tridabase was not made available in DataverseNL. This metadata can only be manually imported. The DCCD portal in DataverseNL makes it possible to store, share and register research data online, during the research period and afterwards. It is also possible to add new data by specialist to the digital archive. Central storage in DataverseNL makes the results of heritage-related tree ring research widely accessible, also for follow-up research. Together with the RCE, DANS made sure that the data from the technically outdated DCCD portal is reusable for future research via this new system.

In the coming months work will be done to realize the harvesting of the DCCD tree-ring data by the ARIADNEplus portal to make integration of this scientific data with other international archaeological data possible.

5 Providing guidelines and support on repository quality control

Task 3.4 will provide guidelines and support partners willing to accredit their repository and their data according to the most important accreditation systems presently leading to CoreTrustSeal, as well as to other systems in use in different EU countries. It will attentively follow the evolution of policies on the matter and provide indications from the archaeological research perspective to international initiatives in the field, for example concerning restriction of access for security and privacy reasons, issues related to language use (multilinguality), or, on the other hand, the implication and an explanation of accreditation requirements when applied to repositories of archaeological data.

DANS-KNAW leads the task, with the support of UoY-ADS, MIBACT-ICCU, DGPC, and SND.

DANS helps heritage institutions that want to start with the certification of their digital archive. By obtaining a certification seal, the organisation makes clear that the sustainability and the reliability of the digital archive is guaranteed. This increases the trust of external stakeholders such as funding agencies, data depositors and data consumers as well as it increases the reputation and visibility of the archive. CoreTrustSeal certification is the most widely used standard for trusted digital archives worldwide. In March 2021, together with the Dutch Digital Heritage Network (NDE) eight recommendations were extracted based on practical experiences of institutions that went through all of the steps of the certification process of CoreTrustSeal (CTS)[24]:

1. Start by reading the CTS Requirements and Extended Guidance. These documents provide a good overview of what might be expected per Requirement, including tips for documentation. Continue by zooming out from this detailed matter and consider for each Requirement how this can be applied to the practice of your digital archive.
2. Do not only collect evidence, but also ensure that these are published and accessible to everyone on your own website. It is better to have a clear and well-arranged story on a part of your website then to supply a huge list of documents.
3. Involve people from across the organisation to gather the information you need. Going through a certification process is a cross-departmental activity. A small team with specific knowledge can write the text for the submission.
4. Assume a reviewer knows nothing about your institute. Explain the context, organisation and infrastructure in a way that can also be understood by outsiders.
5. Usually you don't get the certificate in one go, so don't see this as a failure and do expectation management at the start of the process. In terms of planning, take into account several revision rounds which can take several months.
6. Look for inspiring examples of national and international institutes that have fulfilled certain requirements. All approved CTS applications are publicly available on the CTS website.
7. If you included 'Outsource partners' in your application, make it clear which requirements are fulfilled by the external partner and how this is recorded by contracts (e.g. a service level agreement).
8. Realize that perfection is not the highest goal of a CTS certification process, but that there is a compliance level per Requirement. If things are still in the implementation phase, explain why this is the case and describe the work towards full implementation.

[24] https://wegwijzercertificering.nl/nl/8_tips_uit_de_praktijk/8_tips_uit_de_praktijk

Entering into a certification process gives a new perspective on existing work processes and documentation. This enables the improvement of the organisation itself, which is another important motivation to accredit the repository. Knowing that the policies and practices are 'fit for purpose' gives confidence to the data archive itself.

Organisations which already obtained the CoreTrustSeal have to renew the certification after three years. DANS is one of the first organisations worldwide with a renewed certification. For the renewal of the quality mark, DANS completed the self-assessment which is published online by CTS. However, the work doesn't stop there.

The e-depot for Dutch archaeology embedded within DANS started to become old fashioned in terms of technology on which the EASY archive is running. Over the past 14 years, DANS has provided a designated national online archive to archaeological researchers and the broad public according to formal quality standards. The decision was made to use a new archiving system. Like the DCCD, the archaeological data archive will be migrated to an archiving system based on Dataverse software. Setting up this new Data Station Archaeology, a repository to serve the specific needs of the Dutch archaeological community, is a huge operation. The migration of the data, the mapping of metadata schemas, the implementation of discipline-specific terms and thesauri is part of this process and will improve the findability of specific information, also in regard to the ARIADNEplus portal.

Official assessment guidelines and protocols are part of the certification system archaeological organizations are obliged to follow according to the Dutch Heritage act. It is prescribed that information must be permanently stored in a certified e-depot for durable storage of digital data. On the basis of these provisions, digital files must comply with the preferred formats which DANS defined for each type of file. These are file formats that are internationally trusted to offer the best guarantees for usability, accessibility and sustainability for research in the long term. As part of the evaluation of Cultural Heritage law, working digital instead of analogue in Archaeology becomes the standard practice and the preferred formats of DANS are part of this obligation.

Another major change took place by improving the accessibility of the data in the DANS repository. Datasets only accessible to archaeologists or only visible after a permission request were turned into open access. Thousands of datasets were migrated after consulting and getting approval from the organisations who deposited the data. New tools were created to change the accessibility and the user license to a formal and open access CC-by license from over 45.000 datasets: a remarkable effort, showing the implementation of new policies. Now over 110,000 datasets are available and 99% of these reports, surveys, artefact descriptions and larger excavation archives are openly accessible in the trusted repository DANS offers by having the CoreTrustSeal certificate. It is no longer necessary for users to create an account and to log-in before data can be downloaded.

The launch of this new infrastructure of domain-specific Data Stations at DANS, of which the Data Station Archaeology is one, requests a new self-assessment round of the CoreTrustSeal Requirements to guarantee long-term and secure storage of the archaeological collection according to the newest standards. This will be the focus of next year.

6 Managing FAIRness of archaeological data and IPR

Task 3.5 is led by MiBACT-ICCU and will assess the impact of European and National regulations on data policies in archaeology, with the aim to make archaeological data compliant with FAIR principles. The archaeological data treatment related to IPR issues are complex both for the multiplicity of contents that can be produced and for the presence of the different actors involved. In pursuing this objective, the task analyses the different restrictions imposed by legal regimes and the way archaeologists respond to them, how they are interpreted and how they influence behaviour. Moreover, it identifies good practices specifically in relation to the legal protection of personal data, the protection of intellectual property rights and the use of licenses or waivers to indicate the terms of re-use.

A survey has been carried out in a collaboration of researchers responsible for tasks in two project work packages: Task 3.5 Managing FAIRness of archaeological data and IPR (Flavia Massara, ICCU) and Task 2.2 Reviewing the community needs and the market (Guntram Geser, SRFG). The survey takes account of the task brief, the experience of SRFG from previous ARIADNEplus surveys, and recent other surveys.

Survey context and approach

The context of the survey is determined by the goals of ARIADNEplus and its "sister project" SEADDA to foster the development of archaeological data repositories and enable data discovery and access across existing and newly built repositories. The main purposes of the online survey were to collect and analyse information for assessing the current policies that determine access to and reuse of data held by digital archaeological repositories, and providing guidance and support needed to make the repositories and data FAIR.

ARIADNEplus and SEADDA support principles of FAIR and open data that are being adopted by ever more research funders in their data-related policies. However, working to improve the management and sharing of archaeological data, the projects take full account of the realities of existing practices in the area of archaeology. These practices are in general reasonable when considering various factors such as established rules and routines, limited resources, existing IPR/copyrights, legacy technology and metadata, among others. The objective is not to impose some abstract criteria to fulfil to become "FAIRer" but to support changes in engrained data-related practices so that the outcomes gradually align with the request of being FAIR and open data.

The survey represents a bottom-up approach by focusing on the actual data policies and practices of digital archaeological repositories. This allows an evaluation to what extent these conform to ideals of FAIR and open access data. A reality check in this regard can enable heritage and research authorities, councils and other institutions to reinforce or put in place regulations that bring current repository policies and practices closer to the envisaged ideals. The survey results show that there is quite some room for improvement in this regard.

Survey implementation and results

Survey participants addressed

The questionnaire-based online survey addressed directors, managers and curators of digital repositories for archaeological data. Included in the list of contacts were repositories of ARIADNEplus and SEADDA partners, other known repositories as well as others identified during the survey preparation and dissemination. On the development of the list of survey contacts see below.

Broad definition of an archaeological repository

The survey invited digital archaeological repositories that are operative or in development, in the announcement broadly defined as "*any systems that store and provide access to results from archaeological work in digital formats"*. Concerning repositories currently being set up, the survey announcement explained that in this case *"the answers will concern the envisaged future operation of the repository"*.

Stated purposes of the survey

The repositories were invited to help ARIADNEplus *"assess the current policies concerning access to and re-use of archaeological data"*, and to inform *"guidance on approaches to make archaeological data FAIR (Findable, Accessible, Interoperable and Reusable)"*.

Questionnaire development and testing

The online questionnaire comprises 26 questions, many with several answer options and a free text field for further information and comments. The questionnaire has been implemented on the Microsoft Forms platform and tested by colleagues who work at repositories that are operative or currently being set up. Suggestions for improvements have been implemented.

List of survey contacts

In order to reach and motivate many repositories to fill the questionnaire, ICCU and SRFG created a list of repository contacts. Included were all repository contacts of ARIADNEplus and SEADDA, other known repositories as well as others identified during the survey preparation and dissemination. Some not previously known repositories were identified in online searches with a focus on countries with less coverage on the list, including different regions of countries such as Belgium, the *Länder* in Germany, and the autonomous regions in Spain. Searches looked for repositories of (archaeological) heritage management institutions as well as of research institutes/centres. In addition, registries of repositories have been mined, including OpenDOAR – subject: "history and archaeology", re3data – "ancient cultures", and ROAR – "archaeology" and "history of civilization". Registered university-based and other repositories often use the mentioned subjects among several other subjects to indicate that they have some relevant content. However, these multi-domain repositories have little such content (e.g. some theses, articles, presentations, etc.) and even less, if any, archaeological data. Therefore only few relevant repositories could be added to the list. The final list contained 94 repository contacts.

Survey duration and dissemination

The online survey was open for responses from 17 June to 19 September 2021. In this period ICCU e-mailed directly and individually all 94 contacts, many more than once. The survey has also been disseminated to all ARIADNEplus and SEADDA partners via their Basecamp team communication channel, asking for further dissemination beyond the partnerships. Four contacts said that their organisation does not have a digital archaeological repository. Some other contacts suggested that another person at their institution or a supporting organisation could answer the questionnaire.

Survey responses

The survey gathered information about 60 repositories, 43 operative and 17 currently being set up. Only few respondents did not answer all survey questions. For seven repositories two respondents each provided information. In these cases the data of the questionnaire answered in greater detail was included in the analysis, but where available further information or comments by the second respondent added.

Anonymisation

The respondents were ensured their information would be treated in a confidential manner. Therefore, some information of responses in open text fields have been anonymised where the information allows to identify the institution of the respondent.

Response rate

In total 94 contacts directly received the invitation to participate in the survey. For 60 repositories questionnaires have been filled. These include two from respondents which may or may not have been directly invited to participate in the survey. Taking the 94 directly invited contacts as the basis the survey had a response rate of 64%.

Representativeness

There is no comprehensive overview of repositories that qualify as digital archaeological repositories. Therefore it is not possible to say whether or not the coverage of the survey is representative. However, the survey is to our knowledge the largest thus far on repositories supporting one discipline, as well as the specific topics such as domain data policies and FAIR data. With rich information about 60 repositories the survey results provide insights that further investigations can build on.

Survey participation

The survey gathered information about 60 repositories, 43 operative and 17 currently being set up. The responses provide information on one or more repositories located in most European countries as well as repositories in other countries. Most of the organisations at which the repositories are or will be based are research centres or institutes (20), universities (13), and heritage authorities or agencies (16). The sample of repositories also includes five based at museums, two at archival institutions, and one is being provided by a national archaeological association. Most survey respondents are responsible for more than one task, often including project management, collections development, and digital archiving/curation. 20 of the respondents are directors or deputy-directors of repositories, of which five are also digital archivists/curators. Less present are respondents responsible for IT systems management or user access services and support.

Table 1: Number of repositories per country present in the survey; N=60				
Countries	***Repositories***		***Countries***	***Repositories***
European countries			Netherlands	1
Austria	3		Poland	3
Belgium	2		Portugal	4
Bosnia & Herzegovina	2		Romania	2
Bulgaria	2		Serbia	1
Croatia	2		Slovakia	2
Cyprus	1		Slovenia	1
Czechia	1		Spain	2
Denmark	1		Sweden	2
Estonia	1		Switzerland	2
Finland	1		United Kingdom	2
France	1		**Other countries**	
Germany	3		Argentina	1
Greece	3		Canada	1
Hungary	1		Israel	2
Italy	3		Japan	1
Latvia	1		Turkey	1
Lithuania	2		United States	1
Malta	1			**60**

Table 1 Number of repositories per country present in the survey

Data deposition and curation

Survey questions on data deposition and curation concerned what archaeological work is or will be deposited in the repositories, time after completion of the work until deposition, charge for deposition, embargo period, personal data protection, and long-term storage and preservation.

What is or will be deposited

Most of the repositories contain or will contain results of academic research projects (47) and heritage management work (34) and/or preventive archaeology (30). In a closer analysis, 10 will contain all three categories, 15 academic research and heritage management, and 9 academic research and preventive archaeology. Only 13 respondents did not select the category academic research projects.

This does not mean that most repositories are primarily academic repositories, rather that repositories of both research institutions and heritage management institutions store results of different archaeological work. However, less present in the repositories are or will be results of public or community archaeology projects (22 mentions) and of work of local societies and amateur groups (6).

Time until deposition

Deposition of data from completed archaeological fieldwork usually takes place after one year (16 respondents) or 2-3 years (16). Concerning depositions within 6 months (15), respondents explained that this is mandatory documentation to be provided to the heritage authority or agency already during the fieldwork. Depositions over 3 years after completion of the fieldwork are less common (13).

Data deposition charge

Only at four of the repositories depositors have to pay a deposit charge for the preservation of their data, and at two of these only in case it exceeds a certain data volume.

Embargo period

At 38 repositories depositors can set an embargo period; in comments respondents mentioned periods between 6 months and 5 years, and one even 10 years.

Personal data protection

Concerning personal data related to or within deposited content, respondents could select from three pre-defined answers and also give others or comments. The pre-defined options were: “Require informed consent by research participants, including consent for data storage and sharing”, “Where needed, protect people’s identities by anonymising data”, and “Consider access restrictions to sensitive data”. 15 respondents said that all three measures are being applied, while 12 indicated only informed consent, 7 only anonymisation, and 10 only access restrictions for sensitive data. Other respondents said that two of the measures are being applied.

Long-term storage and preservation

The respondents were asked whether the repository has its own or an external solution for long-term storage and preservation of archived data. 49 said that they had their own solution in place, 11 that an external solution is being used. Respondents who gave further information described the setup of the solution, including internal and external components (e.g. backup), or all provided externally.

FAIRness and data access policies

Over the last few years, the FAIR data principles, published in 2016, have been adopted by research funders, institutes, and researchers to promote the access to research data through data repositories and infrastructures. However, it cannot be assumed that researchers and repositories have a wide knowledge of how to apply the principles in practice. In the international Figshare “The State of Open Data” surveys the percentage of researchers who claimed being familiar with FAIR increased from 15% in 2018 to 20% in 2020. Other respondents had heard of FAIR, but did not consider themselves familiar with the principles, or had never heard of the principles. Researchers’ awareness of the principles and the understanding of what the principles require in practice must be improved. Surveys on the compliance of repositories with the FAIR principles have shown that their implementation is often not sufficient. Many misconceptions of repositories related to the principles’ definition and implementation have been identified.

FAIR-related questions of the archaeological repositories survey concerned (meta)data identifiers, metadata richness, vocabulary in use, data discovery, and licensing. Questions on FAIR principles that

are very technical (e.g., communication protocols) and some specific metadata related questions (e.g., formal knowledge representation or qualified references to other (meta)data) were avoided.

(Meta)data identifiers

Nearly half of the repositories (29) already assign globally unique and persistent identifiers to deposited data, often DOIs but also Handles or Archival Resource Key (ARK) identifiers.

Metadata richness

The majority of repositories (47) thought that deposited data are described with rich metadata (i.e., many descriptive attributes). Dublin Core and archival metadata standards, i.e., Encoded Archival Description or General International Standard Archival Description (ISAD(G) were mentioned in the replies.

Vocabulary support

The repositories use different kinds of vocabularies concerning the user community (international, national, or only by the repository) and standards (e.g., following thesaurus standards, list of terms or keywords given by depositors). Most of the repositories use more than one vocabulary (39), 20 indicated use of two, 17 three, and 2 up to four. Most often an own standardized vocabulary is being used (35), at nine repositories only one such vocabulary (e.g., thesaurus). Also, a national vocabulary and/or an own list of terms is being used quite often at 25 repositories, but seldom as the only vocabulary. In addition, also an international vocabulary (19) and/or keywords given by depositors (17) are being used, but also seldom as the only vocabulary.

Repository search interface

36 of the repositories said that they provide a metadata search interface, while 24 did not or not yet. Among the latter some may provide other ways to navigate and browse information about their collections.

External search platforms

35 repositories did not share metadata with external search platforms. It appears that some do not see a need to make their holdings findable also via external search platforms or for some other reasons cannot do this. There can be many reasons, for example, the user base of the repository is well known and not expected to increase, lack of a suitable external platform, a legacy metadata management system that does not support metadata harvesting.

Copyrights

The FAIR data principles do not address copyrights but it is important to know who holds copyrights. In this regard two main types of repositories and copyright policies can be distinguished: 26 repositories in our sample are mainly or only for depositors that are external to the organisation, while 34 institutional repositories are mainly or only for works of own staff and affiliated researchers. A large number of organisations hold copyrights in works created by own staff (36), while at others the copyrights are held solely by the researchers (15). Only in six cases the organisation holds the copyrights for some works the researchers and for other works. Some organisations also hold copyrights in commissioned works.

License frameworks

At the repositories different license frameworks are applied. A very restricted approach is present at 19 repositories, e.g., all or most works are fully copyright protected and/or own terms and conditions are applied that include some restrictions (e.g., non-commercial, no derivatives or other). 16 repositories have an open approach, e.g., only Public Domain Dedication, only Attribution, or both. 8 repositories which all hold Public Domain data have a mixed approach, while other data requires setting various restrictions, defined by own terms & conditions or standard licenses. Furthermore, 17

repositories apply various restrictions, 12 their own terms and conditions, including some restrictions. Only 5 do not allow commercial use of content.

Enabling open data access

Survey questions on open data access concerned whether there are policies for such access, restrictions applied by repositories, how to improve data access, and how to demonstrate that data is being reused.

Support of open data policies

Repositories need policies and guidelines so that they can support open data access and reuse. Most needed is a clear position of heritage authorities in this regard: 39 repositories required regulations and 36 clear guidelines by the authorities. Next came the challenge to overcome barriers impeding users to deposit open and reusable data (29) which, for example, includes concerns of researchers about open licensing and that their data might be misused. Respondents also considered just as important the training of repository staff to support new policies on open/FAIR data (28). Some respondents also considered that sharing good practices and appropriate technical systems could greatly help to support open data access and reuse policies.

Regulation of archaeological documentation

The respondents were asked whether there is a national legislation in their country that determines which documentation of archaeological investigations and interventions has to be provided to a repository. 36 respondents said “Yes”, 24 replied “No”. Some commented that there is a regulation but it is perceived as insufficient; a lack of or a not appropriate repository was also often mentioned.

Directive (EU) 2019/1024 on public sector information

Until 16 July 2021, the EU Member States had to transpose the Directive (EU) 2019/1024 on Open Data and the Re-use of Public Sector Information into national law. Article 10 of the Directive aims to make research data funded, collected or generated by public sector bodies openly accessible and re-useable. The Article focuses on their institutional or subject-based repositories. Under the definition of public sector bodies fall governmental heritage authorities at all levels (national/regional/local), heritage agencies or associations established by public law, research-intensive public museums and other heritage institutions. The respondents were asked, *“If your repository is located in the European Union, does it fall under the Directive (EU) 2019/1024?”*. 46 respondents answered the question, of which 21 said “Yes”, 5 “No”, and 20 “Don’t know”.

Control of data access

Answers to the question of how people can access data stored in the repositories showed three different approaches. 24 repositories had an open access approach, i.e., no registration is required. 15 repositories had data that can be accessed without registration and other data accessible for legitimate registered users and/or with permission granted. At 21 repositories data was accessible only for legitimate registered users and/or with permission granted.

Improving data access

Answers to the question on what would help the repositories most for improving data access showed that four options were considered more often than others:

- improve or replace the existing data management system (30 respondents),
- improve the quality of metadata (34),
- provide metadata to external search platforms/engines (27),
- use Linked Data to interlink own and other (meta)data (26).

Analysis of the responses for repositories in preparation (17) showed:

- these often wanted to improve the data management system (11), and more often than others also to align their own vocabulary with others (e.g., international or national thesaurus) and/or use advanced ontologies (e.g., CIDOC-CRM);
- repositories that were satisfied with their data management and vocabulary wanted their data to be found by providing metadata to external search platforms and possibly interlink own and other (meta)data using the Linked Data approach.

In the responses of repositories in operation (43) some indicated common priorities while others did not:

- quite a few repositories wanted to improve or replace the existing data management system (19); for nine of these the main reason appeared to be enabling better access to complex or high-volume data objects (e.g., 3D models, LiDAR data);
- also, two smaller groups of repositories with other common priorities could be distinguished: one group primarily wanted to improve the metadata quality and to replace or align their own with other vocabularies; another group had in common the priorities to provide metadata to external search platforms and possibly interlink own and other (meta)data using the Linked Data approach.

Measuring data access, and access during the COVID-19 pandemic

When asked whether the repository collects and analyses access data, 29 out of 56 respondents said "No", 27 "Yes". The latter respondents were further asked whether there has been an increase or decrease of access during the COVID-19 pandemic. Only three said there was a decrease, while 24 reported increases ranging from 5% to over 100%.

Data re-use is difficult to demonstrate

For repositories, data re-use is difficult to demonstrate because if there is re-reuse it takes place outside of what they can easily track and measure. Indeed, when asked whether the repository collects information about data re-use (e.g., references in publications or other sources), only nine of 56 respondents said "Yes".

Conclusions and suggested actions

Conclusions from the survey results and suggestions for activities of ARIADNEplus, SEADDA and other initiatives:

Repository support of FAIR data

- *(Meta)data identifiers:* 29 of the 60 repositories surveyed assign globally unique and persistent identifiers to deposited data, but for many more this would be beneficial. Initiatives for state-of-the-art repositories should provide advice on how to assign such identifiers.
- *Metadata richness:* The majority of repositories (47) are satisfied with the metadata they provide, which suggests no need for targeted support activities. However, in the responses to the question what would help most for improving data access, 34 repositories considered improvement of the quality of metadata. Hence this is still an important topic for advice on good practice.
- *Vocabulary support:* Most of the repositories use more than one vocabulary (39), often two (20) or three (17). Often an own standardised vocabulary (35) and/or a national vocabulary (25) is being applied. However, quite a few of the repositories use less formalised means such as an own list of terms and/or keywords given by depositors (e.g., eight use only this, nine also use in addition

an own standardised vocabulary). Therefore, advice on how to standardise vocabulary and/or align it with an international one (e.g., Getty AAT) would be beneficial.

- *Data discovery:* 24 repositories do not have a metadata search interface and 35 do not share metadata with external search platforms. The reasons for this would be worth investigating in order to advise on how metadata could be provided to data search platforms such as the ARIADNE portal.

- *License frameworks:* While 16 repositories have a very open approach regarding data re-use, 19 have a very restricted one and 17 repositories apply some restrictions. Advice on copyright clearance or why some restrictions should be reconsidered (e.g., no commercial use, no derivatives or other) may be helpful for increasing the potential of data re-use.

Enabling open data access

- *Support of open data policies:* a clear position of heritage authorities is needed; 39 repositories required regulations and 36 clear guidelines by the authorities. Other support is also needed, for example, 28 repositories considered training of repository staff to support new policies on open/FAIR data just as important.

- *Regulation of archaeological documentation:* 36 respondents said that there is a national legislation in their country that determines which documentation of archaeological investigations and interventions has to be provided to a repository, while 24 said there is no such legislation. Some perceived existing regulations as insufficient and often a lack of or a not appropriate repository was also mentioned. Thus, in many countries regulations for rich archaeological documentation and appropriate repositories for such documentation would be needed.

- *Directive (EU) 2019/1024 on public sector information:* when asked whether their repository is located in the European Union and falls under this Directive, 21 out of 46 respondents said “Yes”, 5 “No”, and 20 “Don’t know”. This suggests that more legal support for repositories is needed to understand whether the Directive also applies to them, and what the consequences are if this is the case.

- *Control of data access:* At 21 repositories data is only accessible for legitimate registered users and/or with permission granted. In addition 15 repositories have such restrictions for some of the data, while 24 repositories have an open access approach (i.e., no registration is required). Reducing the barriers to data access would require mechanisms for not disclosing sensitive data and advice could be given on this topic.

- *Improving data access:* The repositories considered what would help them most for improving data access and often this included:

  - improving or replacing the existing data management system (30 respondents),
  - improving the quality of metadata (34),
  - providing metadata to external search platforms/engines (27),
  - using Linked Data to interlink own and other (meta)data (26).

Analysing separately the responses for repositories in preparation (17) and in operation (43) showed some specific needs. For example:

- Repositories in preparation that were satisfied with their data management and vocabulary wanted their data to be found by providing metadata to external search platforms and possibly interlink own and other (meta)data using the Linked Data approach.
- For some repositories in operation that were not satisfied with their data management system, the main reason appeared to be enabling better access to complex or high-volume data objects (e.g., 3D models, LiDAR data).
- Among the repositories in operation, one group primarily wanted to improve the metadata quality and to replace or align their own with other vocabulary, while another group wanted to provide metadata to external search platforms and possibly interlink own and other (meta)data using the Linked Data approach.

The results show that repositories could benefit greatly from advice and support in several areas, in particular, from the perspective of ARIADNEplus, regarding improvement of metadata, providing metadata to the ARIADNE catalogue and portal, and Linked Data.

Analysis of data access and re-use

Repositories also need advice and possibly support regarding collection and analysis of information about data access and re-use:

- *Data access:* 29 out of 56 respondents said that their repository does not collect and analyse data access figures, although this might allow identifying where access procedures and reporting on repository usage could be improved.
- *Data re-use:* No information about data re-use (e.g., references in publications and other sources) is being collected according to 47 of the 56 respondents, although re-use for new research and other purposes best demonstrates that funds for data preservation and access are well invested.

Finally, 24 of the 27 repositories which analyse data access reported that during the COVID-19 pandemic overall there was an increased access, reporting increases from 5% to over 100%, and this is encouraging for the open/FAIR data agenda. It seems likely that the COVID-19 crisis made archaeologists more aware of the importance of publicly shared data, data repositories and discovery and access services.

7 Training on FAIR Data Management

Task 3.6 is led by DANS-KNAW and will organize workshops on FAIR data management. After the workshop in year 2 (as presented in D3.1), the next one is planned for year 4.

The Training Hub was set up as an alternative to workshops and seminars which became impossible to hold in person from March 2020 due to COVID-19. The training topics were based upon feedback from the User Needs Survey, the Hub being launched at the end of Period 1 with updates during Period 2. One of the nine topics is "Applying open/the Fair Principles to archaeology" and this section contains the following five resources:

1. *The FAIR Principles* – online module from PARTHENOS which provides a comprehensive overview using case studies and videos along with the texts.
2. *FAIR Principles and Trusted Repositories* - how DANS combines and operationalizes the FAIR principles and repository certification, a method for assessing FAIRness of data in a video of a presentation by Peter Doorn.
3. *FAIR Data in Trustworthy repositories: the basics* - this video illustrates how certified digital repositories contribute to making and keeping research data findable, accessible, interoperable and reusable (FAIR). Marjan Grootveld of DANS gives a presentation on behalf of OpenAIRE.
4. *PARTHENOS Guidelines to FAIRify data management and make data reusable* – a compact set of 20 guidelines to align the efforts of data producers, data archivists and data users in humanities and social sciences to align research data to the FAIR Principles originally developed in the PARTHENOS Project (in English).
5. *FAIR-Aware* is an online tool developed by FAIRsFAIR which is designed to helps researchers and data managers assess how much they know about the requirements for making datasets FAIR before uploading them into a data repository.

During the last period, three further translations of the *Guidelines to FAIRify data management and make data reusable* have been produced in Czech, Portuguese and Turkish and uploaded to Zenodo. These have proved to be very popular with viewing and download figures (as of 30th November 2021) as follows:

- ***Turkish***: "Veri Yönetimi ve verinin yeniden kullanımı için FAIR Prensipleri Rehberi" (https://doi.org/10.5281/zenodo.3937149)
    - Views: 552 and Downloads: 359

- ***Czech:*** "ZÁSADY zajištení FAIRové správy a využitelnosti dat" (https://doi.org/10.5281/zenodo.3946100)
    - Views: 393 and Downloads: 256

- ***Portuguese***: "Diretrizes para aplicação dos princípios FAIR à gestão e reutilização de dados" (https://doi.org/10.5281/zenodo.3937183)
    - Views: 478 and Downloads: 327

In collaboration with the University of Madrid and the Spanish Association for Digital Humanities (Humanidades Digitales Hispánicas/HDH) the translation of the FAIR Guidelines into Spanish is in sight, and also archaeological partners from Sweden are interested to provide the guidelines in their own language. The Guidelines prove to be an excellent starting point for different stakeholders in each country to understand and help each other in making and keeping data FAIR in a joint journey.

8 Conclusions

The teams working on the different tasks of Work Package 3 Policies and Good Practices for FAIR Data Management continued their work in setting up good practices, guidelines, policy support tools, and in dissemination and training activities on FAIR Data Management in Trustworthy Data Repositories keeping in mind national and international regulations.

After this interim phase, the final period will produce a complete overview of the new insights, guidelines and interactive knowledge exchange, which will be reported in the final deliverable at Month 48 at December 2022 in D3.3: Final report on policies and strategies.