Advanced Research Infrastructure for Archaeological Dataset Networking in Europe

ARIADNEplus is funded by the European Commission's Horizon 2020 Programme

Table of Contents

Executive Summary

This document summarizes the scope, purpose, and activities of the ARIADNE*plus* project.
The first section, **Project Overview**, explains the project goals, i.e. *extending* the archaeological domains served and the users addressed, and *innovating* research methodology as regards digital archaeology by enabling integrated access to archaeological datasets created in Europe (and beyond) with advanced services for further elaborating them in a cloud environment.
The second section, **Concept and Methodology**, discusses how such results will be achieved. It also includes a description of the planned knowledge organization and the technological means to be established for this purpose, i.e. the description of the data infrastructure and of the services to process such data. A diagram at the end of the section summarizes the production pipeline.
The third section, **Impact**, describes the planned impact of the project on the research community and on society at large; and the means – including dissemination and communication – which will be activated to maximize such impact.
The fourth section, **Implementation**, outlines the activities to be carried out by the project.
The fifth section, **Project Consortium** includes the Consortium composition and the individual contributions by partners.
The report is concluded by a number of Appendices:

- Appendix A contains a list of all the partners.
- Appendix B describes in greater detail how the various archaeological sub-domains will be addressed.
- Appendix C describes the use of the CIDOC CRM ontology and its extensions to be used in the project to manage the data.

1. Project Overview

1.1 Objectives

ARIADNE*plus* builds on the success of ARIADNE (Archaeological Research Infrastructure for Archaeological Data Networking in Europe, http://portal.ariadne-infrastructure.eu/), a project funded by the EC in FP7 for the period 2013-2017. ARIADNE succeeded in integrating archaeological datasets in its Registry, with more than 1,700,000 datasets recorded and managed according to the FAIR principles; building a community of use consisting of about 11,000 archaeologists engaged with its services, corresponding to one third of all European archaeologists and probably more than 50% of those using some computer support in their research, with an even higher percentage if only early career researchers are considered; proposing standards for the sector as CRMarcheo, an extension of the well-known CIDOC CRM ontology created to serve the specific needs of the archaeological community; and developing innovative services. The success of ARIADNE is well summarized in a statement from the ESFRI 2016 Roadmap: "*The enthusiastic reviews of these initiatives testify the success of their action to advance knowledge and to establish a research community, acknowledged as "advanced" in official EU documents concerning conservation, or quickly growing in the field of archaeology as shown by the performance indicators of the relevant project ARIADNE. […] In the archaeological sciences the ARIADNE network developed out of the vital need to develop infrastructures for the management and integration of archaeological data at a European level. As a digital infrastructure for archaeological research ARIADNE brings together and integrates existing archaeological research data infrastructures so that researchers can use the various distributed datasets and technologies*."

The overall goal of ARIADNE*plus*, may be summarized as: "*Extending and Focusing ARIADNE*".

Extending concerns the domains served and the users addressed; it has several dimensions, such as:

- The **geographic coverage**, which in ARIADNE already reached almost all the European regions, by integrating in the ARIADNE*plus* infrastructure a greater number of archaeological partners with a particular attention to areas where the coverage was less intensive;
- The **disciplinary coverage**, which in ARIADNE included mainly excavation data and a few other topics as, for example, dendrochronology, by integrating in the new ARIADNE*plus* Data Infrastructure data produced by palaeoanthropology, bioarchaeology, environmental archaeology as well as the results of scientific analyses, such as material sciences, dating and so on, and those related to standing structures, be they small remains of ancient constructions or complex and massive monuments as, for example, Hadrian's Wall in the UK or the Magna Graecia temples in Southern Italy;
- The **time-span** considered, pushing back the earliest datasets included, by incorporating palaeoanthropology, and forwards the end-date until recent times, e.g. including industrial archaeology and Cold War archaeology; in practice, covering the full timespan of the human presence on Earth;
- The **depth of database integration**, exploiting the potential of well-structured datasets such as databases, for which the interoperability will be extended to item level (in ARIADNE implemented only experimentally), and archaeological Geographic Information Systems (GIS), for which integration will be achieved through the introduction of dedicated services, going beyond mere digital maps and overcoming incompatible reference systems;
- The **integration of text datasets** by extending the use of Text Mining through Natural Language Processing (NLP) and Named Entity Recognition (NER), previously applied only experimentally;
- The **research community** involved. The ARIADNE*plus* target is to make contact with the majority of all researchers and professionals (particularly important in this domain where research and heritage management often go hand in hand), and address most if not all the needs of the computer-aware archaeologists. From already existing manifestations of interest, among others in the USA, Japan, Latin America and Australia, it is also anticipated that ARIADNE*plus* will attract the international research community.
- The **service portfolio** offered to users, incorporating more advanced tools for digital analysis and interpretation in the ARIADNE*plus* System.

Focusing has, instead, only one object: **innovation**. This 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 European Open Science Cloud (EOSC): the ARC, Archaeological Research Cloud. ARIADNE*plus* 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).
If innovation in digital archaeology is the Polar Star of ARIADNE*plus*, the project will be guided by a constellation of principles and strategies, such as:

- The Open Science strategy, incorporating Open Access, as fostered by the European Union;
- The Open Data EU strategy;

- The FAIR data principles to make data Findable, Accessible, Interoperable and Re-usable, consolidating what was already applied in ARIADNE;
- Principles for trusted digital repositories and related accreditation, such as the CoreTrustSeal;
- RDA (Research Data Alliance) recommendations, such as those concerning Data Citation and Permanent Identifiers (PID).

Last, but not least, ARIADNE*plus* has a close connection with the initiatives related to the **European Year of Cultural Heritage**. Integrating heritage documentation, and enabling access to it, as planned in ARIADNE*plus* is a way to extend the beneficial effects of this important EU initiative in the future.

What we expect to achieve at the end of the project, i.e. our mission, 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.*

1.2 Ambition

The completion of the ARIADNE project fostered a new approach to archaeological research. Data that before were often considered a mere support to documentation started to become the support for new investigations. Important national experiences like those of the Archaeology Data Service (ADS) in the UK had already generated the birth of similar initiatives in other countries, but they still remained fragmented and unable to cross modern administrative borders. With ARIADNE, such experience became the foundations on which Europe-wide integration was built and discovery across boundaries was made possible. This stimulated the growth of a new attitude towards cooperative research, fostering the concepts of data sharing and re-use and a potential new methodology in which data were one of the pillars of archaeological investigations. With ARIADNE*plus* this possibility will become operational. Extending the integration of data infrastructures to all European regions, all archaeological sectors and all human presence, as ARIADNE*plus* will do, will enable all researchers to make use of the powerful tools made available by the project not only to discover what others have already found, but also to aggregate, link and process such results to create new knowledge. It will introduce into archaeological methodology the concept of "data-based research" that has brought so many advantages in other disciplines. This is unthinkable at present, because the amount of born-digital content and of references to earlier non-digital content is in continuous growth, and can be managed only with appropriate information technology. In this regard, the timelag faced by archaeology becomes an advantage as it allows the use of innovative technology without being constrained by existing ones. With ARIADNE*plus*, archaeology will be able to participate in the EOSC as a protagonist.

The Networking Activities planned in ARIADNE*plus* will support large scale adoption of this new perspective. The project will carry out activities aimed at increasing the number of "followers" from the current 10.000 users of the ARIADNE discovery service to a much larger audience. The collaboration with associations and institutional bodies will facilitate the community penetration of the ARIADNE*plus* approach. Researchers will be encouraged to pursue data FAIRness for the availability of guidelines and tools provided by ARIADNE*plus* and tailored to the specific needs of this research community. Researchers willing to participate in the integration process with their data will find a established and safe procedure as most of the issues potentially raising in the integration process will have been tackled, covering almost every possible subdomain or application. In this way ARIADNE*plus* will achieve not only a critical mass of data, but also a critical mass of supporters. Being a prerequisite for adopting the ARIADNE*plus* philosophy of data sharing and re-use, the culture of cooperation will be fostered not only among participants but also in the wider research community. The starting point for this cultural leap is Virtual Access to an unprecedented mass of data. Researcher participation will be supported by training activities, with Transnational Access being at the top with the provision of personalized hands-on training to develop integration processes with user's own content. This novel approach to archaeological research is supported by standardization, state-of-art technology and innovative services, fit for the needs of 21st-century archaeologists.

The standardization promoted by ARIADNE*plus* covers all the archaeological domains and allows interoperability. The data infrastructure developed by our Joint Research Activities will be supported by cloud technology and will enable the creation of Linked Data. The planned services will enable the data processing that archaeological research demands: first of all discovery, of course, within a huge amount of datasets provided by the integrated data infrastructures; furthermore, services to manipulate the results and produce visual content and annotation tools. Text mining is then a much needed functionality in a field where much data consists of textual reports. Furthermore, the multilingual approach, supported by multilingual vocabularies, is of paramount importance. Finally, being able to process data according to space and time and obtain synthetic views of the results is a key factor to stimulate imagination and lead to new knowledge.

The novel approach to archaeological research described above is one of the innovation aspects that the project will introduce in the research community. The innovative services developed by ARIADNE*plus* and the novel data organization it proposes for integration are other innovation factors. Altogether they are the starting point for groundbreaking innovation: bringing the digital into a discipline that has made materiality its distinctive character.

2 Concept and Methodology

Concept: the archaeological perspective

The purpose of this section is to explain how ARIADNE$_{plus}$ will achieve the objective of global integration of archaeological data infrastructures in Europe, extending it to fields that were not (or only partially) addressed by ARIADNE. It also describes how the underlying technology will be improved and updated, in line with the most recent strategies of the EU, enabling a deeper data integration and creating new environments for data-centred research and data reuse, through the standardization, openness and interoperability of existing key infrastructures. Furthermore, the creation of new innovative infrastructures will be fostered where they do not yet exist, ready to be integrated in the overarching ARIADNE$_{plus}$ framework. The section also outlines how the problems of achieving such ambitious objectives will be overcome through the different actions undertaken by the project.
The underlying concept may be summarized in one simple motto: to make archaeological data FAIR (Findable – Interoperable – Accessible – Reusable) the archaeological data integrating infrastructure must be ILSOFT, as explained in the following table and detailed below.

For archaeological data to be	The integrating infrastructure must be
Findable	**Inclusive**: all key data sources must participate in the integrating effort and effective search tools must be in place
Accessible	**Linked**: source data must be directly linked to the infrastructure catalogue in an integrated/distributed system
Interoperable	**Standardized**: the infrastructure must provide metadata in a homogenous way **Operational**: the underlying technology must provide cross-datasets processing functionalities
Re-usable	**Functional**: services tailored to archaeological research questions must be available **Trustworthy**: data suitability to research questions, data quality and provenance information must be guaranteed by rich metadata

a) Inclusiveness

2.1 The three dimensions of archaeological data infrastructures

Archaeological data infrastructures may concern different aspects of archaeological research. Usually they cover a delimited region, as discussed in the next section. They may address a limited time span: although archaeology in principle covers the time from the appearance of first hominids to the present, data infrastructures may include data pertaining to a part only of the Quaternary period. Finally, although archaeological sciences produce a huge amount of digital data, in general only the conclusions of analytical experiments have been used in archaeological research and included in data infrastructures. Such attitudes to disregard the importance of source data are now being phased out by a trend attributing them with a greater importance. Thus ARIADNE$_{plus}$ will extend its integration process to data that received less attention in ARIADNE, extending both its temporal scope and its thematic scope. The following diagram (Fig. 1) qualitatively illustrates this integration expansion.

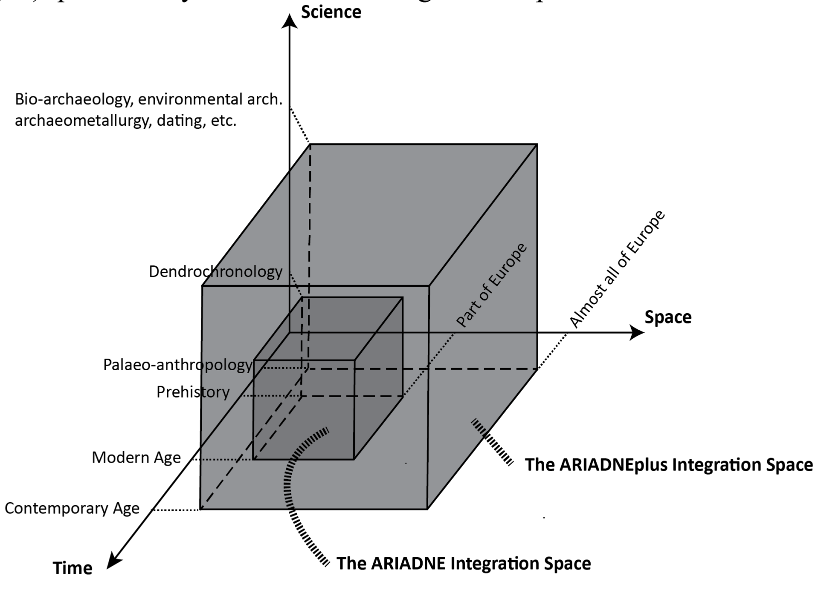

Fig. 1. How ARIADNE$_{plus}$ will extend its integration scope in time, space and content

2.2 Geographic extension of archaeological key data infrastructures integrated into ARIADNE*plus*

As a general rule, general-purpose data repositories focus on archaeological work in the region where the repository is placed, or on archaeological assets (e.g. museum collections, monuments and sites) located in the same region. Sometimes, a much smaller part of their data comes from the work of archaeological missions abroad.

This is true also for thematic repositories, e.g. those containing data from scientific analyses, where some integrating initiatives have already been undertaken. In these cases, data refer to contiguous regions, for example Nordic regions, parts of the Mediterranean basin and so on. In sum, it may be stated that archaeological data infrastructures are largely location-based; their content is limited to modern countries, usually without crossing their borders, or their administrative sub-regions, where cultural heritage and archaeology is in charge of federal administrative institutions, like Germany and Spain. Without integration initiatives like ARIADNE, this situation is clearly a serious obstacle to data-based research, as archaeological research questions concern regions of the past that overlap modern countries and cross modern political borders. Easy examples are the Bell-Beaker culture, spanning in various European regions during the Bronze Age; or the Roman provinces forming Gallia, corresponding to present-time France, as well as parts of Spain, Belgium, Switzerland, Germany and the Netherlands; and so on. Due to variability in time, no geographical rearrangement of data may suit all the archaeological needs spanning prehistory to recent times.

If ARIADNE did not exist, any archaeological study concerning such research questions would need to access a number of data repositories located in each of the relevant regions/countries, and possibly others of a thematic nature, with no common perspective nor common user interface and general search tools – let alone research services.

Furthermore, the data schemas and languages used vary among the different data infrastructures, according to the diverse national official languages and approaches, so any integration process requires a labour-demanding process of collaborative work with data owners, concerning the mapping of the metadata descriptors of individual datasets or the mapping of data schemas of their infrastructure. In addition, incorporating their language in multilingual thesauri and other authority files, if not already present, is also necessary. Such collaboration requires a significant effort in time and resources, especially for large data infrastructures: for them voluntary adherence of repository managers to the integrating process is in general feasible only if some funding is added to available local resources.

The approach adopted by ARIADNE took account of this requirement. Its action developed on two levels:

- Convincing data owners that sharing their data was socially important and scientifically useful, as the advantages of a global registry where all the archaeological information was recorded and made available would benefit each researcher in the archaeological community.
- Support the effort necessary to start the integration process, providing the resources required by data providers.

When forerunners and early adopters take up this approach and the integration process produces a critical mass, the major problems present in the starting phase are resolved, and the effort required to join in is less demanding, but still present. This first objective was achieved by ARIADNE.

On a different note, the ARIADNE project provided the technology necessary for the integration process, with a task force of technological project partners working on it to set up the structure and functionalities of the ARIADNE Registry. They also collaborated with content-providing partners in the mapping and data ingestion process from their datasets. Different but concurring evidence that the integration process has started to gain traction is the fact that during ARIADNE several other important data infrastructures offered their availability and interest in participating in the integration process, but with a few notable exceptions it was impossible to accomplish such work for lack of resources on the providers' side. In sum, we may conclude that:

- Key archaeological data infrastructures are generally location-based; with a few exceptions for scientific thematic repositories.
- The absence of one or more of such key data infrastructures reduces the effectiveness of all the integration, in proportion to the importance and the extension of the local infrastructure, as they are the sole place where data to be integrated can be found. Large gaps would weaken the integrated infrastructure accordingly.
- Starting the integration process requires resources, which local infrastructure managers do not have available. Maintaining it can instead be supported with much less trouble and with no (or very limited) external funding.

Thus, we can conclude that *most, if not all, national data infrastructures are key ones*, as their absence would reduce the integration power of the ARIADNE initiative and its impact on archaeological research; and that *their managing organizations must be involved in the project as partners* to provide them with the resources necessary to start the integration process and to carry it out completely. The complete list of project partners is provided in Appendix A.

In the former ARIADNE project there were clear gaps that ARIADNE*plus* will address. With ARIADNE*plus*, the coverage of archaeological data infrastructures is substantially improved. Fig. 2 illustrates how the geographic coverage will be extended.

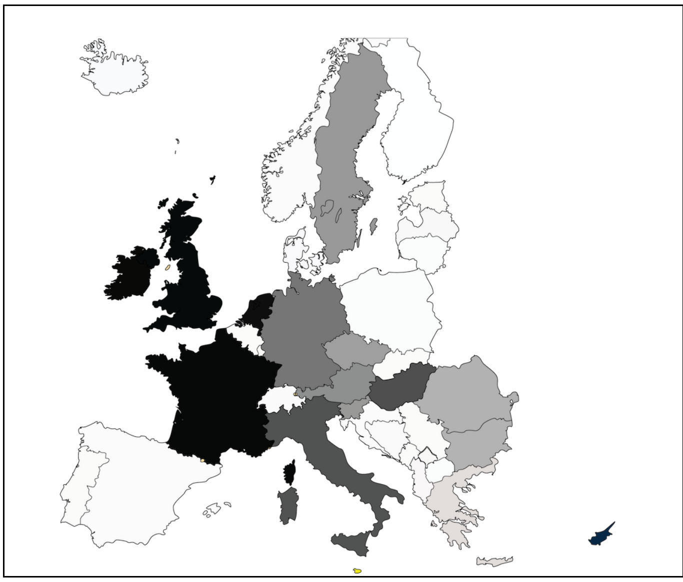

Fig. 2a. Geographic coverage of ARIADNE

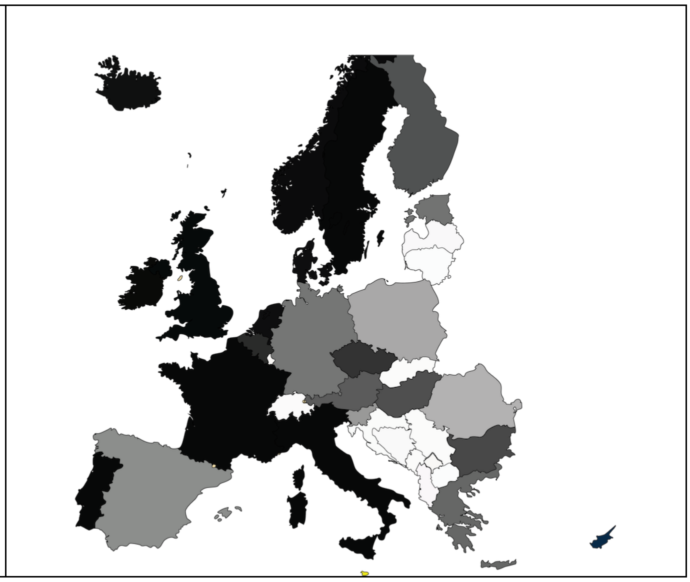

Fig. 2b. Geographic coverage of ARIADNE*plus*

In the above maps, the intensity of the fill colour represents the degree of coverage (poor – partial – good – total) of the online archaeological repositories integrated into ARIADNE. Some countries, especially those where a data infrastructure is missing, are planning to participate thanks to collaboration with the CARARE Europe-wide association. In such countries, ARIADNE*plus* will endeavour to foster and support a culture of archaeological data publication and sharing, through its networking and TNA activities.

With the creation of a European research infrastructure, ARIADNE*plus* will create a RI of global importance. It already attracts and integrates the most important US initiative tDAR, coordinated by the University of Arizona, and the Japanese one created by the national Nara Research Centre, which will contribute their datasets to the ARIADNE Data Infrastructure and promote the ARIADNE*plus* strategy in their respective countries. In Latin America, Argentina is also participating via its national research council CONICET. Beside the addition of their repositories, the most important aspect of this international network is the adoption of the ARIADNE methodology, which is thus becoming globally accepted.

2.3 Thematic extension of the ARIADNE*plus* Data Infrastructure

At first sight, the present ARIADNE catalogue seems to mainly contain datasets on monuments and sites. This is in part true, and in part a misleading effect of the way in which datasets are recorded in the registry. Regardless of its size, a database of finds, for example, is recorded as a single digital object in the registry. This approach is very efficient when the content referred is formed by collections of digital objects, such as the typical archaeological excavation records consisting in texts (the reports), images and drawings (of finds, monuments and sites), and sometimes including also the results of scientific analyses. It is less efficient as regards other databases, where it does not offer any way to directly access individual records or to select choice criteria, as it provides no item level integration. ARIADNE*plus* will address this issue not only by including new artefact datasets, but also enabling item-based searches across them through their federation. Thus an unstructured data repository will be registered in ARIADNE*plus* as a unique digital object, as in ARIADNE. Structured databases, more frequently used for finds and for scientific databases, will instead be federated and access will be enabled across them. As regards unstructured datasets with any content, in ARIADNE*plus* the CRMpe model will be adopted for their metadata organization. This model will be described in greater detail below; here it suffices to state that it has been developed within the PARTHENOS project as a general model for all humanities and heritage catalogues, since it is capable to deal unitarily with activities, actors, procedures, datasets and software. CRMpe is a significant step forward compared with the data model initially used in ARIADNE. The existing ARIADNE catalogue with its 1,700,000 data records has already been converted to this data model and has effectively functioned on this basis for the last year. On the other hand, as concerns the different database schemas, the semantic glue enabling their federation will be the common ontology. In such cases a mapping from the schema used in the database to the CRM or an appropriate extension of it will be defined. A list of the different thematic domains offering databases to be incorporated in the ARIADNE*plus* Data Infrastructure with the above-mentioned process (definition of a common ontology and database federation) follows.

- Human Palaeo-biology and Palaeo-environments
    - Palaeoanthropology
    - 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
- Monuments and Sites
    - Standing Structures
    - Spatio-temporal Data (GIS)
    - Maritime and underwater Archaeology
    - Archaeological Fieldwork
- Inscriptions

These thematic fields are more extensively described in Appendix B.

b) Linked-ness

2.4 The relationship between the ARIADNE*plus* Data Infrastructure and its sources

In the first version of the ARIADNE Catalogue, the relationship with the original sources has been achieved through linking them to the catalogue item referral. Thus users perform searches on the catalogue using the search tools provided, and at the end of the search refinement they obtain a list of results, which are linked to the original dataset(s). This choice avoided the creation of a monster centralised archive and preserves the content provider's control on the data they produce and own. The same search approach will be supported in ARIADNE*plus*. The landing page also enabled user identification when required; in most cases it was not necessary, so anonymous users could directly access the original resources.

The descriptions collected in the ARIADNE Catalogue were structured according to the ARIADNE Catalogue Data Model (ACDM), which was developed in a consensus-based process involving all stakeholders, who could express their requirements and make sure the ACDM accommodated them. Thanks to this process, the ARIADNE Catalogue was built and used as a basis for the discovery service of the ARIADNE Portal. Later in the ARIADNE project, the ACDM was mapped onto the CRM, and the mappings were used, along with a IRI (Internationalized Resource Identifier, extending upon the existing URI - Uniform Resources Identifier) generation policy, to generate the ARIADNE Linked Open Data Cloud. The transformation of the ARIADNE Catalogue into a LOD dataset achieved a twofold effect: from one side, it increased the technical accessibility of the Catalogue, by making it available to the external world via a SPARQL endpoint. From the other, it increased the semantic interoperability of the Catalogue by upgrading it from a domain- and purpose-specific vocabulary (the ACDM) to a general ontology standardized by ISO. ARIADNE*plus* will build on this experience, by adopting the ARIADNE LOD cloud as the core of its information base, called the ARIADNE*plus* data and knowledge Cloud (AC for short). The AC will be born as a LOD dataset, as semantic technologies have reached a mature state of development and, as a consequence, the required resources (ontologies and terminologies) have been made available in one of the semantic web languages. Moreover, the ARIADNE*plus* AC will extend the ARIADNE Catalogue along several dimensions:

1. the ontological dimension: the AC will add to the data resources represented in the ARIADNE Catalogue, also representations of the infrastructural entities that participate in the ARIADNE*plus* network, such as people, institutions and services. Moreover it will provide a finer-grained characterization of all entities, inspired by recent developments in the modelling of infrastructural resources, such as those of the PARTHENOS Entity model. Notice that this extension is already linked to the CRM ontology;
2. the datatype dimension: the AC will add to the types of data represented in the ARIADNE Catalogue a whole range of new data types, as illustrated in detail in the previous sections;
3. the institutional dimension: the AC will add to the data of the ARIADNE Catalogue, the descriptions provided by the new partners in the ARIADNE*plus* consortium;
4. the granularity dimension: the AC will include the data resulting from the item-level integration of some datasets collected from the ARIADNE*plus* network, including overlapping data that can be used to address specific use cases. This extension will push forward the item-level integration experiment successfully conducted by ARIADNE on coin data coming from different institutional repositories and archives.

The richness of the ARIADNE*plus* AC will be exploited in a more sophisticated semantic discovery service offered via its portal and in a much larger LOD dataset offered in access to the whole archaeological research community.

Concept: the knowledge organization perspective

c) Standardization and Operability

2.5 Standardization

In ARIADNE the infrastructure was built around a standard that is progressively becoming universally accepted in the archaeological domain: the CIDOC Conceptual Reference Model (CRM). (http://www.cidoc-crm.org/). The CRM has rapidly conquered a large portion of the Cultural Heritage domain in the large, as witnessed by its influence on the Europeana Data Model. CIDOC CRM has created a number of extensions suitable for use within ARIADNE*plus*, better described in Appendix C.
ARIADNE*plus* standardization will be defined bottom-up.
It will comprise the following steps:

1. Definition of the necessary application profiles through the collaboration of selected users, i.e. archaeologists with expertise in the different domains providing data for integration, and experts in knowledge engineering. This will produce draft proposals to be checked against perceived user needs and then formalized.
2. Mappings are created (with the mapping service X3ML) and tested on samples from the various involved datasets. Feedback will suggest amendments. This will produce the necessary mapping sets.
3. Vocabularies and gazetteers are selected and improved/adapted.
4. When the mapping is stable, the mapping/conversion/ingestion process starts.

This process will be reviewed at mid-term, to improve the fit of the mapping.

Concept: the technological perspective

2.6 Operations enabled by ARIADNE*plus*

User operations in the ARIADNE*plus* framework are enabled by the underlying technological infrastructure and by the services that rely on it. Before illustrating both in detail, we give below an outline of some use cases.

- Search and browse by an anonymous user
  The search function will probably be the most requested one. It will be usable by anybody simply accessing the Portal and availing of this functionality. Search will work in a way much similar as it does on the present ARIADNE Portal, but will be much improved by the greater richness of the ARIADNE*plus* AC, as explained above. Moreover, the availability of multilingual vocabularies will allow searching the ARIADNE*plus* AC in various languages although the data remain in their original language and the metadata are in English. The search is based on the search service activated in the ARIADNE*plus* system. Browsing will be supported by organizing the ARIADNE*plus* data space according to the semantic categories supported by the ARIADNE Portal (time, space and topic) extended as described above along the ontological, institutional, datatype and granularity dimensions. Navigation within the space will be permitted by the very LOD nature of the ARIADNE*plus* AC contents, so that it will be possible to use the resource identifiers as access points to the descriptions of the referenced entities.
- Search and browse by a registered user
  This does not substantially differ from searching and browsing by an anonymous user, with two important differences:
  - It enables direct access to source data stored in repositories requiring registration, if they support ARIADNE*plus* identity federation;
  - It enables storing the search results in a private area within the system and use them for further processing in a Virtual Research Environment (VRE).
- Data processing by an anonymous user
  There is a limited number of services not requiring registration. This is the case of visual services, for example. In general, such services will be those that do not require temporary storage in the ARIADNE*plus* system. When storing is instead necessary, it is impossible to use the service without knowing who is doing it.
- Data processing by a registered user
  Registered users will own their allocation of space in the ARIADNE*plus* Cloud, which will enable the creation of a Virtual Research Environment, shared in total or in part with other researchers if the owner so wishes. This is not just a common (or private) storage space: it will allow activating services on the data stored there and depositing there the processing results.

Such operations will be enabled by the technological level of infrastructure integration achieved by ARIADNE*plus*. ARIADNE adopted an approach for the integration of key infrastructures that avoided to move data around Europe and concentrate them into a unique gigantic deposit. Such a solution would have been hard and expensive to implement, difficult or impossible to maintain and update, and probably ineffective as far as facilitating research is concerned. Instead, ARIADNE created a registry based on the description of datasets stored in the participating infrastructures according to a common data model. The registry can be accessed through the ARIADNE portal and searched by keyword, or according to facets such as place (where), time (when) and content (what). The place and time facets avail of a graphical interface; all three are based on multilingual thesauri. Scientific data were not systematically addressed in ARIADNE, with some notable exceptions as, for example, dendrochronology.

Nevertheless, ARIADNE was able to create a registry with about 1,700,000 records of great utility for data-based research in archaeology. The goal of ARIADNEplus is not only extending the ARIADNE registry by incorporating new infrastructures in it, as explained in the previous sections; but also to improve the level of integration. This will be achieved by several means:

- Adopting an improved data model for the registry.
- Supporting item-based integration for databases, through database federation.
- Exploiting GIS-based data through item-based integration and services for spatial and temporal analysis.
- Linking text reports to each other and to databases via a common ontology, through NLP (natural Language Processing) and NER (Named Entity Recognition) applied to texts.
- Enhance the role of scientific data linking them functionally to their archaeological context and across investigations for comparisons and syntheses.
- Creating a cloud-based environment where all the above functionalities are implemented, with additional services made available to researchers for data-based investigations.

2.7 The ARIADNEplus Cloud and the ARIADNEplus Data Infrastructure

The ARIADNEplus Data Infrastructure will be offered on top of the D4Science infrastructure (d4science.org) and will consist of a hardware layer and a service layer. The hardware layer will be organized as a dynamic pool of virtual machines, supporting computation and storage, while the services layer will be organized into e-infrastructure middleware, storage, and end user services. The hardware layer will consist of an OpenStack installation, supporting the deployment of services in the upper layer by provision of computational and storage resources. The service layer, illustrated in Figure 5 below, will consist of five service frameworks, which can be summarized as follows:

- Enabling Framework: the enabling framework includes services required to support the operation of all services and the VREs supported by such services. As such it includes: a resource registry service, to which all e-infrastructure resources (data sources, services, computational nodes, etc.) can be dynamically (de)registered and discovered by user and other services; Authentication and Authorization services, as well as Accounting Services, capable of both granting and tracking access and usage actions from users; and a VRE manager, capable of deploying in the collaborative framework VREs inclusive of a selected number of “applications”, generally intended as sets of interacting services.
- Storage Framework: the storage framework includes services for efficient, advanced, and on-demand management of digital data, encoded as: files in a distributed file system, collection of metadata records, and time series in spatial databases; such services are used by all other services in the architecture, except the enabling framework; in particular, the ARIADNEplus AC will be implemented on top of the Storage Framework.
- Information Cloud Framework: the information cloud framework includes all services required to collect, harmonize (transformation), and provide (indexing in different formats and backend) all metadata records describing objects, and links between them, of interest to the PARTHENOS community; the ARIADNEplus Data Infrastructure will be implemented on top of the Information Cloud Framework.
- Analytics Framework: the analytics framework includes the services required for running methods provided by scientists taking advantage, in transparent way, of the power of the underlying computation cloud (e.g. parallel computation) and of a plethora of standard statistics methods, provided out of the box to compute over given input data; the Analytics Framework will be used to account for the Virtual Access to the ARIADNEplus infrastructure offered by the project. Moreover, it will be used in the evaluation of the project impacts, offered to text mining and NLP services and to the ARIADNEplus pilots.
- Collaborative framework: the collaborative framework includes all Virtual Research Environments (VREs) deployed by the scientists and for each of them provides social networking services, user management services, shared workspace services, and WebUI access to the information cloud and to the analytics framework, via analytics laboratory services. ARIADNEplus will use VREs to support the management of the project, as well as the services (including the ARIADNEplus Portal) and the pilots developed by the JRA of the project.

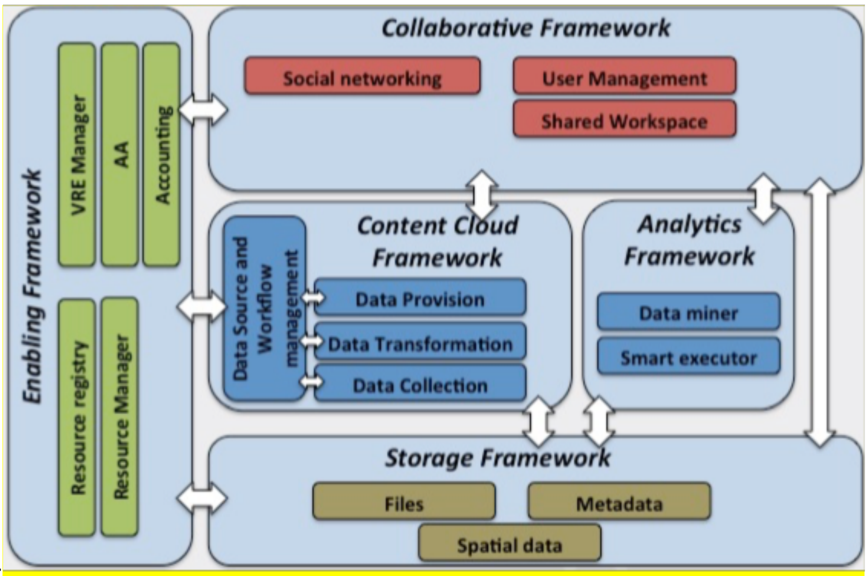

Fig. 3. The service layer of the ARIADNE*plus* Infrastructure

2.8 The ARIADNE*plus* services

Services implemented within ARIADNE*plus* may be grouped as front-office services, i.e. those accessed by users, and back-office services, i.e. those used internally to perform other operations.
Front-office services have been chosen as the ones most popular when using data in archaeological research.

- ARIADNE*plus* visual services
  - Visualization of archaeological imagery. It enhances the visual service already available in ARIADNE and builds on the VisualMedia EOSCpilot Science Demonstrator (http://eoscpilot.eu/science- demonstrators). It enables to display visual archaeological information (images, 3D models) in a fast and efficient way. VisualMedia is already developed in a cloud framework, and will be adapted to the specific archaeological needs, as well as made available in a way coherent with the ARIADNE*plus* service interface.
  - Visual organization of archaeological data. It builds on a tool already developed by CNR to link archaeological documentation to the 3D model of an artefact or monument and visualize it accordingly. Besides a general revision, the existing tool needs porting in the ARIADNE*plus* cloud environment and adapting it to the service interface.
  - Visual documentation of an archaeological excavation. The service builds on the *Ephemera* service (http://ephemera.cyi.ac.cy) used to visualize in 3D the layers of archaeological excavations and the related documentation. Besides a general revision, it will need porting in the ARIADNE*plus* cloud environment and adapting it to the service interface.
- ARIADNE*plus* annotation services
  - Archaeological text annotation. The service consists in an annotation tool for archaeological reports and, in general, texts concerning archaeology. It will build on Open Source annotation tools as, for example, Pundit. The service will be supported by custom archaeological, multilingual vocabularies (also used for NLP and queries).
  - Image annotation. This service is an extension and implementation in the ARIADNE*plus* cloud framework of the annotation tool DAP (http://tss.isti.cnr.it/dap) developed in the TSS project. The tool allows annotating archaeological images in a CIDOC CRM compliant way.
- ARIADNE*plus* text mining and NLP services
  - ARIADNE*plus* NLP/NER (Named Entity Recognition) service. The service is based on the previous ARIADNE text mining tool, further developed into TEXTCROWD, a cloud-based NLP tool created as a Science Demonstrator within the EOSCpilot EU project (http://eoscpilot.eu/science- demonstrators). Work will consist in porting the previous ARIADNE NLP (Natural Language Processing) tool in the cloud environment, following what has already been done for TEXTCROWD, and extending the NLP functionality to other languages beyond Italian (as done in TEXTCROWD), English and Dutch (as done in the ARIADNE tool). Also this task will make extensive use of the vocabularies, gazetteers and time period vocabularies.
- ARIADNE*plus* space-time services
  - GIS services. They consist in the usual services present in GIS systems, for example buffer definition, layer selection, proximity, viewshed analysis and so on, and will be supported by D4Science Cloud geoserver

GeoTools, which already has many of the required functionalities built-in. It will also rely on the gazetteers and named time periods vocabularies developed within the project.

Back-office services are those required in the background to support other services or to enable machine access to the data. They have been already mentioned in the previous sections, here we just recall the ones more directly related to users:

- The ARIADNE$_{plus}$ query services include the necessary functions supporting queries on the ARIADNE$_{plus}$ catalog. It will also include a multilingual query capability based on the multilingual vocabularies.
- The Cloud geoserver will implement the usual GIS functionalities used by the GIS service.

d) Functionality and Trustworthiness

2.9 Archaeological data-centric research questions and related services

It has been argued that archaeologists should excavate less in favour of better exploiting the content already amassed and stored. This is increasingly feasible as the quantity of digital data about new discoveries, or converting/referring to previous non-digital results, is growing. A new Big Data paradigm is appearing in archaeology and in all digital humanities: not the datasets with peta- and perhaps exabytes of numbers, typical of Nuclear Physics, but millions of small files characterize the so-called "long tail of science", where archaeology belongs with the distinctive characteristics of combining in its knowledge base texts, quantities and visual content Addressing archaeological Grand Challenges such as, for example, early technological advancements, migrations, trade relations, and so on, requires dealing with bits of information dispersed in thousands of datasets hidden among millions of similar (but unrelated) datasets. That is why discovery is the primary concern: creating a sieve retaining only what pertains to a research question. But then, it is still necessary to aggregate, combine and further process those filtered data: this is why processing services are paramount. This is how the term *innovation* translates in the archaeological domain: a new way of combining machine processing with human imagination to creatively address Grand Research Challenges about the past and create new knowledge. Discovery as a service was the primary objective in ARIADNE, and is still so - improved and more flexible - in ARIADNE$_{plus}$. It is complemented by the services described above, which provide a solution to address archaeological research questions. But, above all, ARIADNE$_{plus}$ will create a research ecosystem where data and services are the pillars, openness is the philosophy and collaboration is the attitude.

2.10 Data trustworthiness

On the other hand, the reliability of the results depends on the quality of the data. ARIADNE$_{plus}$ will provide the instruments to support researchers' confidence on data created by others, establishing policies for repository quality such as repository accreditation, and foster the extensive adoption of Data Management Plans by data creators, which will need to document how such data were generated, by whom, with which purpose (it may make a difference if a 3D model, for example, was created for dissemination purposes or for research), and how they were curated and maintained. The compliance to ARIADNE$_{plus}$ recommended policies and good practices, as defined in the related workpackage specifically for the archaeological domain, and the information on their provenance will inform users about their reliability. Not fulfilling the ARIADNE$_{plus}$ recommendations may suggest that care should be taken when re-using such data. The ultimate responsibility stays however with the re-user: caveat emptor.

2.11 Conclusions

The outline of the conceptual approach of ARIADNE$_{plus}$ presented in the previous sections shows that ARIADNE$_{plus}$ is not starting from scratch. It incorporates the excellent results of ARIADNE and builds on them an innovative way of carrying out data-based research in archaeology. It incorporates the results of other EU-funded projects. Among them, we may quote PARTHENOS, which will provide the data model and a number of policy recommendations. Other recent projects providing input are the EOSC (European Open Science Cloud) related ones, such as EOSCpilot and EOSChub. They are also a reference with which ARIADNE$_{plus}$ will endeavour to be compatible. Many of the CRM extension mentioned above were developed within EU projects, for example CRMdig was created in the CASPAR project and refined in 3D-COFORM. D4Science, the backbone of ARIADNE$_{plus}$ processing operations, was originally created in the homonymous project and further developed in D4Science-II, with services - now to be used in ARIADNE$_{plus}$ - which were implemented for the needs and with the support of Research Infrastructure projects in other domains, such as, for example, the geoserver implementation. Extensive usage of Open Source software will be fostered, including the previously mentioned Geo Tools libraries, the libraries used for NLP and many more. In conclusion, ARIADNE$_{plus}$ will not only integrate data infrastructures. It will also integrate and combine ideas, methods, tools, and results of previous research activity, setting up a new framework aimed at innovating the methodology and the practice of archaeological research.

Methodology: technology meets users' needs

The ARIADNE*plus* methodology comprises different components that intersect and are combined to implement the concepts described in the previous section. As shown in fig. 4, teams with different expertise in the ARIADNE*plus* partnership cooperate in the project activities. Extending and supporting the ARIADNE*plus* community addresses both archaeological partners and stakeholders. Such activities aim at consolidating and extending the existing ARIADNE*plus* community. Special attention is paid to Central and Southeastern Europe, with a dedicated subtask. The collection of information aimed at defining community needs, actually an update of previous work done within ARIADNE, will also be undertaken. Already existing liaisons with major associations and international bodies such as, amongst others, EAA, the European Association of Archaeologists, and EAC, the European Archaeological Council (the presidents of both sit on the ARIADNE*plus* Strategic Advisory Board, see section 5) are further reinforced. A task is dedicated to professionals and heritage managers, the "industry" sector of archaeology. Finally, given that grand archaeological research challenges – such as climate change and migration – are not confined to the continent of Europe, international collaborations are cultivated and extended.

Developing policies and strategies for FAIR data management targets both archaeological partners and the community at large. The good practices recommendations developed here and the tools, e.g. a DMP template for archaeological data, are provided to users. The ARIADNE*plus* KOS is developed through the collaboration of domain experts, i.e. archaeological partners, and KOS experts, as already mentioned. The application profiles and the vocabularies produced in this activity are then used to create the mappings from each infrastructure to be integrated in ARIADNE*plus*. All the aspects explained in the previous section are considered for integration. This is preliminary to the ingestion of data/metadata in the ARIADNE*plus* Data Infrastructure (ADI), which updates the former ARIADNE catalogue and has been set up in the meantime by the project IT team. The ADI is hosted in the ARIADNE*plus* Cloud (AC) where also the services developed by the IT team can operate. Such services are defined from the beginning, but their features are defined according to the needs indicated by users. The ADI and the services enable the creation of Virtual Research Environments in the ARIADNE*plus* Cloud, planned to be an integral component of the European Open Science Cloud.

Fig. 4 illustrates the flow and interconnections of activities in ARIADNE*plus*. For the sake of clarity, some activities are not included.

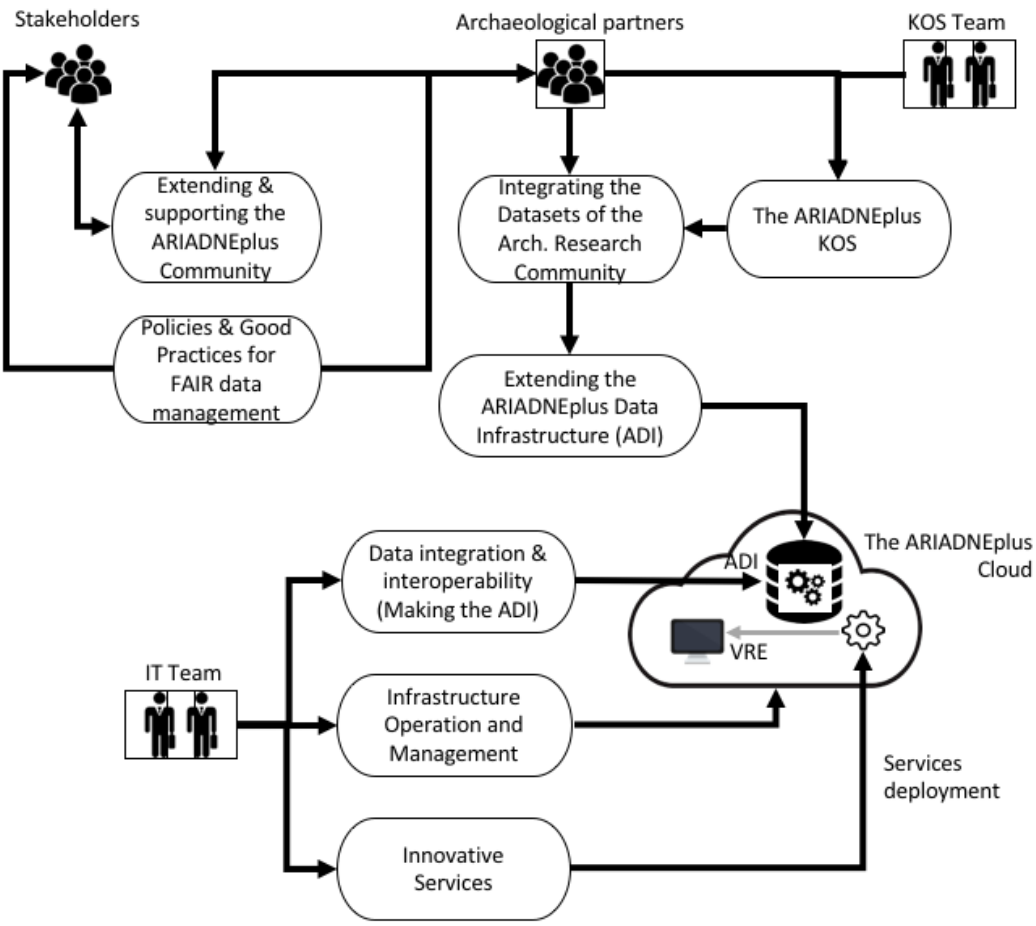

Fig. 4. Work methodology in ARIADNE*plus*

3. Impact

3.1 Expected impacts

This section describes the expected impacts of ARIADNE$_{plus}$ on the archaeological research community and, in general, on society.

[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.
As a provider of discovery, access and other online services ARIADNE$_{plus}$ supports researchers' needs with regard to data and tools irrespective of the location where they conduct their research, thanks to the move of all services to Cloud-based provision implemented as Virtual Research Environments (VREs).

[2] New or more advanced research infrastructure services, enabling leading-edge or multidisciplinary research, will be made available to a wider user community.
More advanced and new research infrastructure services will be implemented with regard to metadata and vocabulary enrichment and integration as well as data access and uses.

[3] Operators of related infrastructures will develop synergies and complementary capabilities, leading to improved and harmonised services. There will be 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.
With regard to research infrastructure services ARIADNE$_{plus}$ will register its portfolio of cloud-based and other web services in the catalogue of the European Open Science Cloud (EOSC), currently in development by the EOSChub and EOSCpilot projects. This will allow other digital research infrastructures using complementary ARIADNE$_{plus}$ services relevant for their operation and *vice versa.*
ARIADNE$_{plus}$ aims for full integration into the EOSC. This will enable, within the EOSC initiative, optimisation of operations (harmonization, cost savings) as well as common development of digital services of providers of related fields (i.e. archaeology, various other humanities, heritage sciences) and relevant others such as environmental and geo-sciences data services.

[4] Innovation will be fostered through a reinforced partnership of research organisations with industry.
Access to more and richer data and information about archaeological sites, monuments and objects will also benefit businesses active in heritage-led development projects (i.e. in urban or regional regeneration) and cultural heritage tourism. Moreover, creative businesses that produce media, i.e. 3D models and other imagery, for museums and information centres at heritage sites (i.e. virtual reconstructions) can benefit especially from using the advanced ARIADNE$_{plus}$ visual media services for the communication of cultural heritage.

[5] A new generation of researchers will be educated that is ready to optimally exploit all the essential tools for their research.
Overall, beneficiaries of ARIADNE$_{plus}$ TNA and training offers will be educated to use tools and services for archaeology and cultural heritage e-science, using virtual research environments (VREs) for advanced collaborative research and research data management.

[6] Closer interactions between larger number of researchers active in and around a number of infrastructures will facilitate cross-disciplinary fertilisations and a wider sharing of information, knowledge and technologies across fields and between academia and industry.
Sharing and linking data through ARIADNE$_{plus}$ various data will stimulate cross-disciplinary fertilisation among researchers in these as well as natural sciences domains. The aim to bring together knowledge and data for these purposes promotes cross-fertilisation with regard to the frameworks and conceptual understanding of different fields of research in which data acquire significance. A strong cross-fertilisation is also being achieved between scholars and developers of software tools for research purposes.

[7] The integration of major scientific equipment or sets of instruments and of knowledge-based resources (collections, archives, structured scientific information, data infrastructures, etc.) will lead to a better management of the continuous flow of data collected or produced by these facilities and resources.
Holders of structured data collections may consider registering these with appropriate metadata and vocabularies in the ARIADNE$_{plus}$ registry, hence use the infrastructure and services to enable discovering and accessing such collections. The ARIADNE$_{plus}$ consortium includes a few partners with relevant data collections such as results from material analyses. Knowledge and experiences from the integration of these collections in the ARIADNE$_{plus}$ infrastructure will be shared by these partners with the research community.

[8] When applicable, the integrated and harmonised access to resources at European level will facilitate the use beyond research and contribute to evidence-based policy making.
We consider that ARIADNE$_{plus}$ can promote interest in questions of archaeology and heritage by citizen scientists and students. The ARIADNE$_{plus}$ portal will provide open and free access to content from providers around Europe and thereby stimulate study work and visits to archaeological museums and sites.

3.2 Measures to maximise impact

a) Dissemination and exploitation of results

[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.
Researchers using the ARIADNE*plus* European infrastructure will have online access to rich knowledge-based data resources from integrated national and domain digital infrastructures, from around Europe and non-European countries (Israel, Japan, Argentina, USA). The ARIADNE*plus* infrastructure allows researchers cross-database search and access to data from different archaeological fields of research.

[2] New or more advanced research infrastructure services, enabling leading-edge or multi-disciplinary research, are made available to a wider user community.
ARIADNE*plus* will make available a range of advanced and new services for research data providers and users, enabling leading-edge and multidisciplinary research. The project will demonstrate advantages of using ARIADNE*plus* services and data to the archaeological research community and other user groups.

[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.
ARIADNE*plus* coordinates its activities with related research infrastructures in the field of humanities and heritage to develop synergies and exploit optimised operations and complementary capabilities. Common development in the field will be continued with regard to harmonised data models (i.e. data catalogues, CIDOC CRM and others). The overall strategy with regard to improved use of resources, economies of scale and cost-savings is Cloud-based virtualisation and integration in the European Open Science Cloud (EOSC).

[4] Innovation is fostered through a reinforced partnership of research organisations with industry.
Private sector actors in archaeology are small businesses of contract archaeologists and consultancies, providing professional services in preventive archaeology. To explore and demonstrate opportunities, project activities address private businesses specifically, including also cultural and creative businesses.

[5] A new generation of researchers is educated that is ready to optimally exploit all the essential tools for their research.
Researchers increasingly need to use advanced tools for data-intensive research paradigms and, at the same time, are requested to acquire skills in data management and sharing based on the FAIR data principles. ARIADNE*plus* supports these requirements through provision of advanced services, Transnational Access to RI competence centres of partners, training offers, and promoting high-quality domain FAIR data and repositories.

[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.

[7] 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.
Holders of structured data collections will register these in the ARIADNE*plus* registry, so that they can be discovered and accessed via the e-infrastructure services.

[8] Integrated and harmonised access to resources at European level can facilitate the use beyond research and contribute to evidence-based policy making.
User groups beyond archaeological and cultural heritage researchers will be addressed where appropriate.

b) Data Management Planning

The project will elaborate, implement and regularly monitor and update a detailed Data Management Plan (DMP), taking full account of the H2020 Guidelines on Open Access and FAIR data management. This brief section conforms to the current reflection of the consortium about the data that will be generated and processed for the purposes of the project. The detailed ARIADNE*plus* DMP for making the data findable, accessible, interoperable and reusable (FAIR) will be provided in Month 6 of the project.

c) Communication activities

The project will prepare and update a systematic plan for communicating its goals, activities and results to its different reference communities. The main objective of the communication is to support the building and networking of a vibrant community of stakeholders in, and users of, the community research infrastructure und services. Communication activities will be planned yearly in a dedicated document.

Communication approach

With 41 partners, the ARIADNE*plus* consortium has a strong communication and networking capacity, with partners being active members of the research and professional communities of different fields of archaeology and cultural heritage. A core element of the communication approach will be highly targeted communication tailored to the different user groups, their perspectives and advantages of using the ARIADNE*plus* infrastructure and services. Communication messages, products and channels will be chosen accordingly. The target communities range from archaeological researchers, interested in scientific progress and the academia, to user groups such as cultural and creative businesses and citizen scientists.

Communication channels and materials

ARIADNE*plus* will make use of a set of communication channels and materials designed following the project branding concept and design patterns, including mention of EU support. The main communication channels and products are:

Project website

A state-of-the-art website will serve as the hub for all ARIADNE*plus* online communication about the project (i.e. news & events) and for downloading the digital versions of information material (i.e. leaflets, presentations, reports).
The website will be professionally designed, pages will be optimized for high impact on search engines, and a complete metadata structure will be present for easy indexing and retrieval. All standard provisions such as information pages about the project context and goals, major themes, services, event calendar, partners, EU funding, contact information, etc. will be implemented also featuring the ARIADNE*plus* research e-infrastructure, services and data resources.

Community channels and social media

The partners and their researchers, data curators and technical experts will involve their network contacts, domain e-mail lists and presence on research and professional platforms to distribute project information and material.
Appropriate external social media platforms will also be used, for example, Twitter for short messaging with links, SlideShare for presentations and other information material, and YouTube for audiovisual content.

Promotion and event material

General and special promotion material: General material will include a project information leaflet, PowerPoint presentation and poster, updated according to project phases. Special material will e.g. include: announcements and leaflets with more specific information to accompany larger project events, memoranda signed with major institutions, the launch of new services, the innovative pilots and other project highlights; for culture and science journalist's material on ARIADNE*plus* themes will include pre-prepared high-quality images and texts.
Event material: A "conference pack", including a roll-up or poster, short project video, project presentation and leaflets, will allow partners to effectively communicate ARIADNE*plus*' goals, activities and results at events.
Basically the ARIADNE promotion materials will be in English; however, templates for local translation and dissemination of material will be available for the project partners.

Conferences

The presence of ARIADNE*plus* at conferences will be planned systematically to achieve an optimal coverage of the different scientific and professional user groups. Events will include the **project kick-off**, with a **public part**, a **mid-term event** and the **final conference**, to present the overall achievements and sustainability plan.
Furthermore, the project will organise 15 sessions, workshops or roundtables at **international and major national conferences** with delegates from other countries.
Project partners that specialise in **particular research domains and/or data** (e.g. LiDAR or 3D models) will present the ARIADNE*plus* approach to their community of peers and "next generation" researchers (the number of participants here ranges between 100 and 300).
A special focus will also be on **digital heritage & archaeology** conferences
Moreover of course conferences on **research e-infrastructures** will receive special attention. We also expect that over the coming years there will be EOSC events at which ARIADNE*plus* will aim to be present.

Publications

Joint and individual publications will be planned to achieve a broad coverage of ARIADNE*plus* results in journals, proceedings or thematic publications. Conference sessions organised by the project, or individual participation, will provide an opportunity for presentations and papers. Digital copies of all scientific publications (preprints or published open access) will be gathered and made accessible on a page of the project website, but also distributed through other effective channels such as SlideShare. Publications of the project such as booklets will be distributed in the same way but also produced in small print-on-demand runs for free distribution at events.

4. Work Plan

The work plan is organized in three groups of activities, plus Management and Control.

The first grouping concerns Networking. One set of activities is in charge of extending and supporting the ARIADNE$_{plus}$ community. It addresses with specific tasks: the partnership, the archaeological research community, the stakeholders and the international community. Another set of activities defines policies and good practices about FAIR data management. Such policies are mainly targeted to researchers and research institutions, but are also relevant for all stakeholders. Finally, a dedicated set of activities addresses infrastructure integration. Application profiles developed by the project are applied to each infrastructure and mappings are defined. Other tasks concern assistance to partners in this operation and integration of ARIADNE$_{plus}$ with existing digital libraries, implemented through the ingestion of metadata.

Knowledge transfer is implemented through the "access" activities where integration projects by partners, other institutions or research teams are supported through visits to a selected group of partners.

Another grouping comprises all technological activities: setup and maintaining the ARIADNE$_{plus}$ Data Infrastructure within the ARIADNE$_{plus}$ Cloud is setup, developing and putting into operations VREs and creating the ARIADNE$_{plus}$ KOS, by defining applications profiles, mappings, vocabularies and gazetteers. Specific activities are in charge of implementing services: visual, annotation, text mining and geo-temporal services as well as back-office services that manage queries. Finally, a number of innovative pilots based on the ARIADNE$_{plus}$ technology and dataset integration are implemented.

5. The Project Consortium

The ARIADNE*plus* Consortium is very broad, but each partner has a precise role in contributing to the project objectives. Since ARIADNE*plus* aims to cover all of Europe and all archaeological domains with its integration process, the large size of the Consortium is a necessity. Partners may be grouped in three categories: those providing content (the "archaeological" partners), those contributing with technological expertise (the "technology" partners) and a couple more addressing policies and innovation (the "strategy" partners). In some cases, expertise is provided via linked third parties. Archaeological partners cover with their infrastructures all the dimensions of ARIADNE*plus* integration: geographical, temporal and disciplinary (e.g. palaeoanthropology, bio-archaeology, environmental archaeology and, in general, all archaeological sciences). To avoid an unmanageable growth of the consortium, key infrastructures have been selected and, in general, no more than one infrastructure per country and/or per thematic domain are involved. Only a couple of exceptions are present, due to the different nature of the partners involved, usually one from a research environment and the other one more oriented to managing large data infrastructures. The table below shows the contribution of each partner to the integration process: "G" denotes a geographic infrastructure, "T" a thematic one.

Partner	Country	G/T	Content contributed
PIN	IT	T	Medieval Italian Churches (buildings), Epigraphs
UoY-ADS	UK	G+T	Sites & monuments, fieldwork, arch. science
OEAW	AT	G	Excavations, site and monuments
BUP	BE	G	Sites & monuments, environmental arch.
NIAM-BAS	BG	G	National archaeological inventory
AMZ	HR	T	Finds and inscriptions
CYI	CY	T	Arch. science (glass), coins
ARUP-CAS	CZ	G	National archaeological inventory
AU	DK	T	Arch. prospection, arch. science (dating, environment)
UH	FI	G	Excavations
INRAP	FR	G	Sites and monuments, preventive archaeology
RGK	DE	T	Finds, coins
ATHENA-RC	EL	G	Excavations
FORTH	EL	T	Bioarchaeology
PP	EL	T	Excavations, prospection
HNM	HU	T	Arch, science (metallurgy, ceramics, dating, environment)
FI	IS	G	National inventory of finds, monuments & sites
CARARE	IE	G	Remote sensing
IAA	IL	G	National inventory of finds, monuments & sites
AIAC	IT	G	Monuments & sites of more than one country
MIBACT-ICCU	IT	G	National GIS of monuments & sites
INFN	IT	T	Arch sciences (metallurgy, ceramics, glass, dating)
UNIROMA1	IT	T	Palaeoanthropology
DANS-KNAW	NL	G+T	Finds, monuments & sites. Dating, environmental arch.
RUG	NL	T	Prospection
KHN-UO	NO	G	Finds, monuments & sites
DGPC	PT	G	National inventory of monuments & sites
LNEC	PT	T	Standing structures, buildings
INP	RO	G	National inventory of monuments & sites
IAVP	RO	T	Finds
ZRC-SAZU	SI	T	Remote sensing
CENIEH	ES	T	Palaeoanthropology
UB	ES	T	Inscriptions
SND	SE	T	Environmental arch., rock carvings, medieval churches

The above table shows that thematic coverage is complete, and the geographic one is almost complete. The presence of the CARARE Association and its Europe-wide network brings into the ARIADNE$_{plus}$ community also countries not yet managing an archaeological data infrastructure.

International partners

Three international partners from Argentina, Japan and USA contribute to the internationalization of the ARIADNE$_{plus}$ approach and to dissemination in their respective regions.

Technological partners

The competence of technological partners covers all the necessary domains.

Partner	Country	Technological expertise
PIN	IT	Semantics
CNR	IT	Semantics, knowledge management, infrastructures and cloud technology, 3D visualization
CNRS	FR	Semantics, 3D data management
ATHENA RC	GR	Digital libraries
FORTH	GR	Knowledge management, semantics
SND	SE	Services, UI
USW	UK	Vocabularies, gazetteers, text mining

Finally, DANS-KNAW addresses policy issues basing on their expertise and participation in European organizations (e.g. ScienceEurope, RDA), while SRFG is in charge of fostering innovation and maximizing impact, building on their experience in this domain.

All the partner institutions are key actors at national level: there are national research centres (CNR, CNRS, ATHENA RC, FORTH, LNEC); ministries and other public organizations (INRAP, FI, IAA, MIBACT-ICCU, DGPC, INP); science academies (OEAW, NIAM-BAS, ARUP-CAS, DANS-KNAW, IAVP, ZEC-SAZU); and top-level academic and research institutions, world leaders in their respective domains.

APPENDIX

Appendix A – ARIADNEplus: the Consortium – List of Partners

Part. No	Participant organisation name	Short name	Country
1 (Coord.)	PIN SOC.CONS. A R.L. - SERVIZI DIDATTICI E SCIENTIFICI PER L'UNIVERSITA DI FIRENZE	PIN	IT
2	UNIVERSITY OF YORK	UoY-ADS	UK
3	CONSIGLIO NAZIONALE DELLE RICERCHE	CNR	IT
4	SALZBURG RESEARCH FORSCHUNGSGESELLSCHAFT M.B.H.	SRFG	AT
5	CONSEJO NACIONAL DE INVESTIGACIONES CIENTIFICAS Y TECNICAS (CONICET)	CONICET	AR
6	OESTERREICHISCHE AKADEMIE DER WISSENSCHAFTEN	OEAW	AT
7	Bruxelles Urbanisme & Patrimoine	BUP	BE
8	NATIONAL INSTITUTE OF ARCHAEOLOGY WITH MUSEUM-BULGARIAN ACADEMY OF SCIENCE	NIAM-BAS	BG
9	Arheoloski muzej u Zagrebu	AMZ	HR
10	THE CYPRUS INSTITUTE	CYI	CY
11	ARCHEOLOGICKY USTAV AV CR PRAHA VVI	ARUP-CAS	CZ
12	AARHUS UNIVERSITET	AU	DK
13	HELSINGIN YLIOPISTO	UH	FI
14	CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS	CNRS	FR
15	INSTITUT NATIONAL DE RECHERCHES ARCHEOLOGIQUES PREVENTIVES	INRAP	FR
16	Roemisch-Germanische Kommission des Deutschen Archaeologischen Instituts	RGK	DE
17	ATHINA-EREVNITIKO KENTRO KAINOTOMIAS STIS TECHNOLOGIES TIS PLIROFOFORIAS, TON EPIKOLNONION KAI TIS GNOSIS	ATHENA-RC	EL
18	FOUNDATION FOR RESEARCH AND TECHNOLOGY HELLAS	FORTH	EL
19	PANEPISTIMIO PATRON	PP	EL
20	MAGYAR NEMZETI MUZEUM	HNM	HU
21	Fornleifastofnun Islands	FI	IS
22	Connecting Archaeology and Architecture in Europe	CARARE	IE
23	Israel Antiquities Authority	IAA	IL
24	ASSOCIAZIONE INTERNAZIONALE DI ARCHEOLOGIA CLASSICA ONLUS	AIAC	IT
25	Istituto Centrale per il Catalogo Unico delle biblioteche italiane e per le informazioni bibliografiche	MIBACT-ICCU	IT
26	ISTITUTO NAZIONALE DI FISICA NUCLEARE	INFN	IT
27	UNIVERSITA DEGLI STUDI DI ROMA LA SAPIENZA	UNIROMA1	IT
28	Nara National Research Institute for Cultural Properties	NARA	JP
29	KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN - KNAW	DANS-KNAW	NL
30	RIJKSUNIVERSITEIT GRONINGEN	RUG	NL
31	UNIVERSITETET I OSLO	KHN-UO	NO
32	DIRECAO GERAL DO PATRIMONIO CULTURAL	DGPC	PT
33	LABORATORIO NACIONAL DE ENGENHARIA CIVIL	LNEC	PT
34	INSTITUTUL NATIONAL AL PATRIMONIULUI	INP	RO
35	INSTITUTUL DE ARHEOLOGIE VASILE PARVAN	IAVP	RO
36	ZNANSTVENORAZISKOVALNI CENTER SLOVENSKE AKADEMIJE ZNANOSTI IN UMETNOSTI	ZRC-SAZU	SI

Part. No	Participant organisation name	Short name	Country
37	CONSORCIO PARA LA CONSTRUCCION, EQUIPAMIENTO Y EXPLOTACION DEL CENTRO NACIONAL DE INVESTIGACION SOBRE LA EVOLUCION HUMANA	CENIEH	ES
38	UNIVERSITAT DE BARCELONA	UB	ES
39	GOETEBORGS UNIVERSITET	SND	SE
40	UNIVERSITY OF SOUTH WALES PRIFYSGOLDE CYMRU	USW	UK
41	Arizona Board of Regents	ASU	US

Appendix B – Archaeological Subdomains and their Relationship with ARIADNEplus

B.1 Scientific data infrastructures and their integration in ARIADNEplus

Archaeological sciences are an important field of archaeological research. They complement with scientific evidence the results of fieldwork, making use of several methods based on physics, chemistry, biology and materials science. The study of palaeo-biology and palaeo-environments also falls under this heading.

Palaeo-anthropology may nowadays use sophisticated forensic methods such as the *analysis of ancient DNA*, combining anthropological data with methods derived from biology. *Biomolecular archaeology*, i.e. mainly but not exclusively the study of ancient DNA, uses technology and techniques developed in other scientific disciplines, so the data organization is derived by these and is usually finalized to process the data with the software tools developed for their purposes. Biological remains are also analysed with specific methods that may involve sophisticated techniques, or just recognition of samples compared to standard specimens, the so-called reference collections. In such cases, for example with pollens, a digital replica may substitute the original reference specimens, and facilitate the remote identification of samples. The importance of data obtained with techniques from Biology and other related disciplines to archaeological research is well demonstrated by the SEAD data infrastructure (https://www.sead.se/) on *environmental archaeology*. This infrastructure has set up a database documenting more than 1600 sites with some 15000 datasets. The SEAD infrastructure will be incorporated in ARIADNEplus.

As concerns archaeological sciences based mostly on Physics, perhaps their most popular deployment is in *dating*. It is evident that the possibility of attaching an absolute date to archaeological remains is a very useful contribution to the understanding of the chronology of past events, in combination with the relative sequence provided by stratigraphy and the interpretation of artefact styles and materials. ARIADNE had already integrated in its registry the data collected for *dendrochronology*, the dating method based on tree-rings. The infrastructure extension will incorporate data related to radiocarbon dating, based on the ^{14}C isotope of carbon as well as other methods such as those based on the relative percentage of chemical elements or their radioactive isotopes, on magnetic properties, and on thermoluminescence. These methods apply to specific materials, for example radiocarbon dating may be used only on organic material. The degree of data standardization among such techniques varies substantially. Data are often unpublished and only the conclusions are included in the archaeological publication.

In a similar way, *materials science* techniques used for example in archaeometallurgy or in the study of ceramics and glass generate a large number of numeric data which are usually presented as spectra.

In conclusion, there are features common to most datasets produced by archaeological sciences:

- With some notable exceptions (e.g. radiocarbon analyses), there are no standards for data management as they are often regarded as disposable after the results are obtained.
- Data are often not published online because they are not perceived as relevant per se, but only instrumental to produce synthetic conclusions. The above mentioned SEAD is a notable exception.
- Important information about the procedure of data acquisition or generation is omitted.
- The link between the analyses and the archaeological context is not fully made explicit and is in general not usable by a machine.
- The lack of easy availability of data online means that analyses may be duplicated by different projects, wasting valuable resources.

These characteristics imply a greater difficulty to organize scientific data. If the knowledge organization system for general archaeological data is well tested – although requiring improvements – there is no such thing for scientific data. Integration of such data will require additional work to define proper extensions of the domain ontology to guarantee interoperability without losing the substantial information produced by the scientific experiment: such extensions must be integrated in the general framework of archaeological data, but also suitable to compare, re-use and re-process the numeric results. There is also a lack of thesauri (let alone multilingual ones) and other authority files that must be filled; in this case, a registry of protocols that can be referred is an example of the additions required. The research community has already started addressing such issues. Further advice from the communities of use will refine and ultimately set up appropriate data models (called CRMsci) within the CRM family.

B.2 Archaeological prospection

This group of archaeological activities concern surveys carried out to collect information prior to excavation. They can be done with no instrumentation, using specific instruments or making use of aerial imagery. The project will address these together, to identify common concepts, and separately, to distinguish the impact of the technology used. In particular, the organization of data coming from metal detector activities will be analysed, which has already generated projects such as Portable Antiquities in the UK and is currently addressed by new initiatives, for example the SUALT network (Finland), DIME (Denmark), MEDEA (Belgium) and PAN (Netherlands). It is however always essential that the activity fulfils all legal requirements, which in some countries require metal detector surveys to obtain a specific permit by the archaeological authorities, usually to prevent looting, illegal trade of antiquities and

destruction of context. Such limitations are a strict application of the Valletta Convention (1992) that states (Art. 3, para ii) “the use of metal detectors … must be subject to specific prior authorization”.
These differences also need to be reflected in data organization, for instance by requiring that the reference to such a permit is recorded. In general, prospection generates data about finds and sites, so data organization is linked to the schemas envisaged in these cases. Finds data also includes databases of objects found in the framework of an organized mission, including field-walking as well as excavation. Dispersed objects, such as those discovered during a small emergency excavation or found occasionally during a survey, may appear in unstructured reports, of which several thousands were integrated in ARIADNE. Finds inventories will be the target of item-level database integration planned by ARIADNE*plus* as one of the main steps forward compared with the current catalogue state.
Archaeological prospection also includes various techniques of remote sensing, including the many types of geophysical survey, and Aerial imagery including LIDAR (Light Detection and Ranging) images, which nowadays provides additional information to the existing practice of using photos taken from above, at various distances, facilitating a global view of existing sites or the discovery of new ones. Such data will be profitably processed with the visual services which ARIADNE*plus* will provide.

B.3 Monuments and sites datasets vs artefact-oriented ones

A large majority of the datasets catalogued in ARIADNE concerned monuments and sites. In this case the extension is straightforward when new datasets are of the same type as the ones previously catalogued in the infrastructure as it envisages only a wider geographic scope. Some particular cases may require the introduction of a different approach.

B.4 Standing structures and buildings archaeology

Although archaeology is usually associated with the idea of excavation – in popular imagination as well as in common practice – archaeological remains also include above-ground structures, still partially or totally standing. Just to quote a few famous examples, archaeology also studies Stonehenge, the Pyramids, the Coliseum and the Athens Acropolis. Together with these iconic monuments, there are many more that need to be studied with a methodology that recognizes the shape and the topology of such structures whether they are 18th- or 19th-century houses which are still occupied today, or if all that survives are just low walls. Documentation of standing structures results from a combination of archaeological and architectural concepts: for example, it includes passages, doors as a way to access one space from another one, and empty spaces delimited by buildings as a courtyard. Systems have been developed, for example in the UK, in Italy and France, for the digital documentation of buildings. The use of Building Information Modelling (BIM) is widely used in the construction industry and is becoming as commonplace in archaeological recording as GIS. However, the use of such systems is not standard throughout Europe, because such documentation systems often mix archaeological remains with historic palaces, and are perceived as not pertaining to archaeological research.
ARIADNE*plus* will start the integration of such data infrastructures, addressing first methodological questions such as how their data must be organized in the project ontology framework, and will test its approach on some selected but important repositories from Italy, UK and others.

B.5 Spatio-temporal data and GIS

Another category of concepts will require preliminary research on how to organize and exploit the information they provide. These are space-time-based data. Of course, time and space are intrinsic components of archaeological knowledge, so dealing properly with them is of paramount importance.
Time management in the current version of ARIADNE keeps into account the fact that archaeological time is almost always imprecise and is often referred using period names rather than time spans based on dates. The time spans corresponding to named periods vary from place to place; there is a famous picture produced by the ARENA project some 15 years ago comparing period names with time spans, as used in six European countries (fig. A1).

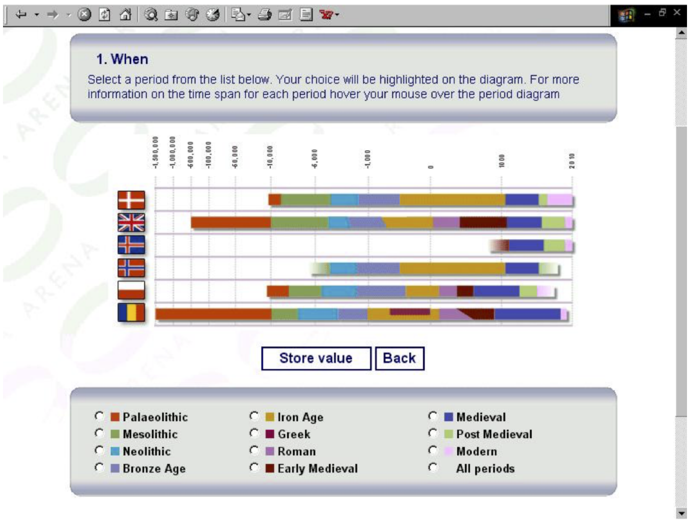

Fig. B1. Correspondence between named periods in six European countries

The difference is not just nominal: it corresponds to differences in the pace of cultural development in different regions, so there is no overall chronology valid throughout Europe, except for modern times. In the ARIADNE registry the issue was managed by asking experts to list named periods used in their country and to assign to each one a time span. However, the granularity of this organization is often too coarse to be precise: firstly because there are denominations specific to particular ambits (e.g. "the Orientalizing Period", a term used by Etruscologists); and then because the geographic area of reference for the correspondence named period – time span may vary in time and may not correspond to a modern country. Furthermore, modifiers of period names are widely used, e.g. "Late Bronze Age" or "Early Roman Period", and in different language variants. In a few cases, different period names co-exist because of different schools or approaches to the same studies, as in the case of Cypriot ceramics where three different chronologies (lists of period names) have been used. As regards period names, ARIADNE*plus* will implement a system to address this complexity, building upon the successful method developed in ARIADNE, based on PeriodO (https://perio.do/), a gazetteer of scholarly definitions of historical and archaeological periods. Also time operations, the so-called Allen algebra, will keep into account the fuzziness of such definitions. This is particularly important for texts and the related NER (Named Entity Recognition) tools. A similar situation affects space. The best currently available gazetteer is Pleiades (https://pleiades.stoa.org/), providing in digital form the location of some 40.000 ancient locations. Pleiades covers the Greek and Roman world, and is extending to Celtic, Byzantine and Early Medieval geography. After that, modern gazetteers such as geonames may be used safely.

The ARIADNE*plus* approach to address archaeological GIS is rather simple. An archaeological GIS consists of an archaeological database with geo-referenced data; a geographic base, describing in a digital way the terrain and its features; and of tools that enable processing the data according to the spatio-temporal concepts. Database views (the result of queries) produce what in GIS terms is called a layer, i.e. a georeferenced selection of the data, for example all those pertaining to a given time period or the sites where a particular kind of pottery was found. Some special tools process data together with the geographic base, for example singling out the data located at a given distance from a given point or line (what in GIS terms is called a "buffer zone"); or those that can be seen from each other (the so-called "viewshed analysis") because there is no geographic feature as a mountain in the middle.

Thus, the ARIADNE*plus* approach will deal with GIS databases integration as it does with any database; it will integrate geographic features across different systems by a simple conversion of their parameters to a Europe-wide system; and will create the necessary services implementing all the required functionalities.

Management of geo-temporal data will be provided by GeoTools, an Open Source Java GIS toolkit already installed and working on the system hosting ARIADNE data, which already has all the necessary functionalities (thematic maps, layers, topology, buffers, and so on) and is easy extendable to incorporate any other additional one thanks to its open source nature and its large community of developers. It will then be possible for example to check if two Iron Age sites, located on facing hills, one in Italy and the other in Slovenia, each one recorded in the respective and separate archaeological GIS, were (and still are!) visible from each other.

The ARIADNE*plus* geoserver will also incorporate an additional functionality that is often ignored in archaeological GIS, taking into account that geographic features may not remain unaltered through time. Assumed stability is acceptable in most circumstances, but there are cases in which it is not true and may be misleading. Rivers, lakes and sea coastlines are typical examples: a riverbed, a river mouth and a coast may change. Ignoring this would locate ancient villages in the middle of nowhere (at present): a good example is the Prile lake in Southern Tuscany, which existed until the Roman Period. It explains the location of Etruscan settlements and Roman villas, once located on its shores but now in the middle of the countryside: in time, this shallow lake was filled by the sediments of a tributary river and nowadays it exists no more. On a similar note, the discovery of two perfectly preserved Roman sea ships

in Pisa, 30 km from the sea, would be a surprise when disregarding the progress of the coastline, again due to the Arno river silt that filled the sea, moving the coast away from the Pisa port. Of course, archaeologists are well aware of the above, but, unless well instructed, a computer program is not.
These considerations and those concerning named periods explain why we speak of spatio-temporal aspects together: dealing separately with time and space concepts is a simplification that is almost always an oversimplification, as these dimensions are always intertwined and as far as possible should be kept so. The selected geoserver enables this approach. ARIADNE*plus* will integrate in this way several GIS already created - or which are in the process of being created - at national level: among others in Norway, Sweden, Italy, the Czech Republic, Hungary, Iceland, Bulgaria, Romania and more.

B.6. Maritime and underwater archaeology.
Archaeological investigations carried out under the surface of lakes and of the sea do not differ, in the basics, from those carried out on the earth surface. There are however a few particular features that call for attention. In some cases (the most famous one is in Baia, near Naples) the remains are under the water surface because of a bradyseism or another natural phenomenon. Intertidal investigations concern the coastal areas uncovered at low tide and covered at high tide. In most other cases, the relevant feature is instead the effect of a shipwreck. For the latter, the global distribution of such events is an important information to interpret complex phenomena such as trade routes, and this relates these investigations to space-time related data considered above. Thus, also for this branch of archaeological research data organization needs to be analyzed.

B.7 Inscriptions
Inscriptions are the source of a large amount of direct and indirect information. We include under this heading proper inscriptions, as analyzed by epigraphy, inscriptions made on objects, such as the ones existing on amphoras that enable to identify their provenance and destination, explored for example by the EPnet project (http://www.roman-ep.net/wb/home/), and rock carvings as collected in the Rock Art project database at U. Gotheborg, Sweden. Both outstanding examples will be integrated in ARIADNE*plus*. Arrangements are being undertaken with EAGLE to incorporate their integrated archives of epigraphs.

B.8 Archaeological fieldwork
Most modern archaeological fieldwork takes the form of what is often known as ”preventive archaeology”, through urban or industrial development, or major transport infrastructure projects. It is the activity that perhaps provides most of the archaeological data concerning the past of modern towns. It provides information about sites of different size, from a Roman villa to an entire town (as in the case of the remains of the Etruscan town discovered at Gonfienti near Prato, Italy, found during the construction of a major logistic centre and truck terminal), and the finds retrieved during such excavations. The data structure may depend upon the recording system used, such as Intrasis in Sweden, or the IADB is some UK projects, and there are issues of interoperability between different standards.

Appendix C – Using CIDOC CRM within ARIADNE*plus*

The CRM is a formal structure (ontology) for describing the concepts used in cultural heritage documentation. Since 2006, it is an ISO standard (ISO 21127). A very important feature is CRM extensibility: it is possible to add concepts required in some specific sub-domain, as a specialization of the model, without losing high-level compatibility with other sub-domains or globally, at the domain level. For example, there is an extension called CRMarcheo for archaeological excavations which is a specialization of the CRM and is compatible – at high level – with any other extension like CRMsci, the specialization for scientific experiments. The adequacy of the CRM to the archaeological domain has been proved by its adoption in ARIADNE as the ontology for the re-modulation of the ACDM and for the item-level integration of the coin data. It must be noted that the CRM allowed the ARIADNE Catalogue and the coin data to be modelled in the same semantic network, allowing users to state queries that is not possible to express otherwise, neither in the context of ARIADNE, neither outside that context. In what follows we highlight the features that make the CRM suitable for the pivot role it is going to play in ARIADNE*plus*.
The CRM allows modelling so-called *perdurants* via the *Temporal Entity* class and *endurants* via the *Persistent Entity* class. In addition, it allows the spatio-temporal contextualization of events (a major temporal entity class) by explicitly modelling Time-Spans, Places and Space-time Volume, when the spatial or temporal projections of events are not known. Furthermore the CRM offers properties to specify the people, the objects or the concepts related to an event, providing a typing mechanism for specifying the specific role played by these entities in the context of that event. Overall, the CRM provides 80+ classes and 120+ properties forming the basic ontology. All those basic concepts are – or may be – further specialized according to need. Specializing subclasses limits interoperability; but this usually happens at the deepest levels of the specialization tree, so a high level of compatibility is in general guaranteed. The CRM SIG, the Special Interest Group that continuously revises and upgrades the CRM ontology to meet the new challenges, in the course of the last decade has created a number of extensions of the CRM, to address the specific representational needs of general domains. Each such extension is an ontology of its own, whose classes and properties are specializations of the classes and properties of the core ontology. The CRM extensions most relevant for ARIADNE*plus* are concisely summarized below:

- CRMarcheo (http://www.cidoc-crm.org/crmarchaeo/) is the CRM extension for archaeological excavation. It introduces all the entities required to document the excavation process, from the class for *Archaeological Excavation* to those classes related to the stratigraphic method as *Stratigraphic Volume Unit* and *Stratigraphic Interface*. The new properties document the stratigraphy generation and features, and the human activity to investigate it.
- CRMsci (http://www.cidoc-crm.org/crmsci/) is the CRM extension defined to deal with and to document generic scientific activity. It incorporates the CRM Entities and Properties, including new concepts such as *Observation*, *Data Evaluation*, *Sample* and more, with the related new properties.
- CRMdig (http://www.cidoc-crm.org/crmdig/) is the CRM extension addressing the digital world, both methods of production and artefacts produced. It introduces concepts such as *Digital Object*, *Digital Device*, *Software* and so on, and properties related to actions typical of digital processing.
- CRMba (http://www.cidoc-crm.org/crmba/) deals with standing structures (buildings archaeology, hence the acronym). It introduces the concept of *Built Work* and various entities related to spaces defined by walls. Moreover, it considers topological relationships such as *is connected to*.
- CRMgeo (http://www.cidoc-crm.org/crmgeo/) addresses geo-temporal issues. It has introduced several concepts, some of which have already been subsumed in the general CRM model, such as Space-time Volume mentioned above, and the related properties such as *occupied*.
- CRMtex (http://www.cidoc-crm.org/crmtex/) finally is the newly born CRM extension dedicated to documenting inscriptions for properties. It introduces concepts such as *Written Text* and *Transcription*, typical of these studies.

From our experience in ARIADNE and in similar work already done by the partners involved in the construction of the ARIADNE Knowledge Organization System (KOS), we expect that most of the concepts necessary for the project implementation are already present in the above-mentioned extensions. In sum, it will not be necessary to construct a new ontology from scratch. It will be sufficient to choose the necessary entities and properties among the already existing ones, all compatible with each other due to CRM extensibility, and perhaps add a few domain-related specifications. Last but not least it is necessary to recall here CRMpe (from Parthenos Entities), the model created by the PARTHENOS project (http://www.parthenos-project.eu/) as the metadata schema for catalogues in the Humanities and Cultural Heritage, which is used in the current ARIADNE catalogue and will also be used in ARADNE*plus*. The main characteristics of CRMpe have already been outlined in the previous section. The main advantage of CRMpe is the capacity of describing, in a unified way, humans and their activities, actions performed by computers and the related procedures, and artefacts resulting from computer or human activities, the latter broadly encompassing all tangible and intangible heritage. A mapping will be established between CRMpe and the metadata model for catalogues defined by the EOSC to guarantee interoperability: the latter is presently (March 2018) under construction, with a preliminary draft just presented. Thus we will achieve full compatibility with the forthcoming EOSC catalogue. The CRMpe is described in the PARTHENOS D5.2 Deliverable, http://www.parthenos-project.eu/Download/Deliverables/D5.2_Report_on_ design_ Joint_Resource_Registry.pdf.

An important aspect of project standardization concerns multilingual vocabularies and gazetteers. The importance of appropriate gazetteers and time-period vocabularies has already been described above. Multilingual vocabularies are no less important. For general archaeological activity, ARIADNE had already used Getty's AAT (http://www.getty.edu/research/tools/vocabularies/aat/about.html). Additions will concern, for example, the scientific domains involved in archaeological sciences, and extensions for more specific finds categories. A specific Task will address this topic.