ARIADNE Report Summary

Project ID: 313193
Funded under: FP7-INFRASTRUCTURES
Country: Italy

Final Report Summary - ARIADNE (Advanced Research Infrastructure for Archaeological Dataset Networking in Europe)

Executive Summary:
The mission of the ARIADNE project (Advanced Research Infrastructure for Archaeological Dataset Networking in Europe) is to bring together and integrate existing archaeological research data infrastructures, so that researchers can use the various distributed datasets and new technologies as an integral part of the archaeological research methodology. To achieve these objectives ARIADNE implemented an integration infrastructure which provides useful information services for archaeology, and related stakeholders, but also to a wider range of potential users requiring access to collections and datasets. Integrating datasets from researchers across Europe into a common portal faced technical challenges. Achieving interoperability, to enable a search across all the different datasets to return relevant results, implicated agreeing on common standards and terminology. Creating a network of researchers and institutions who are interested in sharing their data and in contributing to building standards is the social dimension of ARIADNE was also another important achievement of the project.

From the beginning the project focused on raising awareness about the ARIADNE research infrastructure amongst stakeholders within the partner organisations; research institutions; managers, senior researchers, scholars, researchers and students; international networks; research infrastructures in related disciplines; policy makers and policy bodies; funding agencies including the European Commission; and the public at large.

During its duration the project:
• Identified the main channels for communication and networking with ARIADNE’s stakeholder community including social media, conferences, mailing lists, etc.;
• Built a contact database;
• Participated in clustering activities with research infrastructures and related projects;
• Actively participated in national, international and domain events;
• Shared and exchanged news and information with the stakeholder community about project results, events, training and trans-national access opportunities;
• Developed a set of dissemination materials including the project website, brochures, posters and other materials.

The ARIADNE Portal (http://portal.ariadne-infrastructure.eu) offers a central point of access to over 2 million records of archaeological resources made available from partner institutions throughout Europe. For the first time major collections published by the Archaeology Data Service, DANS, FASTI-Online and the Deutsches Archäologisches Institut in their online catalogues can be searched together along with collections provided by ArheoVest Association, Institute of Archaeology of the Czech Academy of Sciences, the Italian Ministry of Cultural Heritage, the “Athena” Research and Innovation Center, Cyprus Institute, the Discovery Programme, INRAP, Bulgarian Academy of Sciences, Österreichische Akademie der Wissenschaften, Swedish National Data Service and the Slovenian Academy of Sciences and Art.

Project Context and Objectives:
The aim of the project is to overcome the fragmentation and isolation of archaeological datasets in Europe, fostering their interoperability and accessibility for research across institutional and national as well as disciplinary boundaries. The project is based on a combination of three activities, namely Networking and community building, Research and technical development,
and Trans-national access to research centres and online services, all supporting the overall goal mentioned. In particular, the Networking and community building fosters and supports the involvement and participation of institutional stakeholders as well as individual researchers in the project activities. Cooperation agreements have been signed with major institutions across Europe (and beyond) and thousands of researchers and practitioners have been reached and involved through user needs & requirements surveys, project sessions and individual presentations at international and national sector conferences, technical and data management tutorials, participation in meetings and workshops of other projects, and so forth.

Research and technological development concerns all activities required to enable advanced data interoperability, integration and access. This includes data models and vocabularies, e.g. the Data Catalogue Model, CIDOC-CRM extended for archaeology, database and vocabulary mapping tools for linked data. Furthermore, it includes the building of the e-infrastructure and services, including dataset registration, ingestion and integration, and portal and other services. Transnational access (TNA) activities comprise the online access to data services, first provided by individual and then integrated services through the ARIADNE portal, and physical access of researchers to centres of expertise of the ARIADNE consortium. The TNA activities consist of a mixture of research and networking: they support researchers in creating their own archaeological datasets according to good practices and in view of inclusion in the ARIADNE registry and interoperability with other datasets.

Project Results:
The ARIADNE Infrastructure
The goal of ARIADNE, that is to bring together and integrate disperse archaeological resources have been achieved with the implementation of the infrastructure, which aims to provide useful services and tools to the archaeological community. The ARIADNE infrastructure integrates data and metadata from different providers into one common schema (ACDM), and also provides semantic integration based on subject, space, time. This integration intends to provide useful and user-friendly information services for archaeology, and related stakeholders, but also to a wider range of potential users requiring access to collections and datasets. In ARIADNE the aggregation service consists of a set of tools and services that aim to harvest content related to archaeology, and to clean and transform it to a common format. Aggregation includes a number of steps, such as validation, cleaning and enrichment. According to the content aggregation workflow the MORe aggregator was used to harvest existing content in other schemas and map it to ACDM, or import directly from batches of XML or Excel files. ACDM encoded information was then pushed to an RDF store and to the Registry service. Another service was used to synchronize content from the RDF store to an Elasticsearch index service. This process provides researchers with the ability to use heterogeneous distributed datasets as an integral component of archaeological research methodologies.

The infrastructure aggregates over 2 million records, cleaned, enriched and published to the ARIADNE portal (http://portal.ariadneinfrastructure.eu) which is available and used globally.

The portal consists of three main services: Resource discovery, Data integration services, Other services. The resource discovery services rely directly on the registry, whereas the data integration services provide more intelligent ways of data integration through LOD and NLP technologies or through the ARIADNE CRM (Conceptual Reference Model). The solution of mapping ACDM into CRM, an ISO standard widely used within Cultural Heritage, has proved to be an appropriate technical decision, making the ARIADNE Catalogue more interoperable. This interoperability has been exploited in the context of the ARIADNE LOD experiment on the integration of the catalog-level data with the item-level data, for instance in the coin demonstrator.

Archaeological extensions of CIDOC CRM ontology
In order to address the complexity of archaeological data integration, partners within ARIADNE worked on implementing the requirements to employ and extend the CIDOC CRM. The idea of developing archaeological extensions of the CIDOC CRM ontology to cover the specific needs of the archaeological investigation and to help archaeologists examine and understand the complex relationships between all entities and activities related to it, derived from this need. The two archaeological extensions of the CIDOC CRM, CRMba thought to model the complexity of a built structure from the perspective of buildings archaeology, and the CRMarchaeo developed to model the processes involved in the investigation of
subsurface archaeological deposits, were implemented in the framework of the project to provide tools for the exploration of the available archaeological datasets archives and the semantic integration of the digital resources made available within the network.

The research activities carried out during the four years led to the definition of the ARIADNE Reference Model which comprises the following extensions:

- **CRMinf**: the Argumentation Model intended to be used as a global schema for integrating metadata about argumentation and inference making in descriptive and empirical sciences.
- **CRMsci**: the Scientific Observation Model intended to be used as a global schema for integrating metadata about scientific observation, measurements and processed data in descriptive and empirical sciences.
- **CRMgeo**: the spatiotemporal model that provides a linkage between the standards of the geospatial and the Cultural Heritage community in particular between GeoSPARQL and CIDOC CRM.
- **CRMdig**: the model for provenance metadata to encode metadata about the steps and methods of production (“provenance”) of digitization products and synthetic digital representations such as 2D, 3D or even animated models created by various technologies.
- **CRMba**: the Buildings Archaeology ontology developed for investigating historic and prehistoric buildings, the relations between building components, functional spaces, topological relations and construction phases through time and space.
- **CRMarchaeo**: the Excavation Model ontology to encode metadata about the archaeological excavation process.

**Linked Data Cloud**

The ARIADNE registry holds metadata of data resources from the content providers. These metadata are being collected and enriched with the MoRe aggregator and included in the ARIADNE data catalogue. ARIADNE makes the catalogue and other data generated in demonstrators available as Linked Open Data (LOD); thereby the ARIADNE LOD can become part of a web of Linked Data of archaeological and related other information resources.

This work within ARIADNE involved the use of a suitable RDF store and graph database for the Linked Data generation and linking efforts. The project has experimented with Virtuoso and Blazegraph to perform archaeologically relevant SPARQL queries on the generated Linked Data, and to allow updates of datasets using the SPARQL 1.1 Graph Store HTTP Protocol. Based on this work, a scalable implementation that can efficiently support the publication and use of the ARIADNE LOD has been designed and realized to offer three different services: the Linked Open Data Server, the Demonstrators, and the Mapping and Ontology Server. The Linked Open Data Server provides access to a large RDF dataset, which comprises of several graphs of archaeological datasets and can be queried via a SPARQL endpoint. The Demonstrators have been developed to exemplify the capability of Linked Data based item-level data integration to support answering archaeological research questions. A number of datasets about coins, sculptures and wooden material have been integrated based on mappings to the CIDOC CRM (and the archaeological extensions) and use of other domain vocabularies. The Mapping and Ontology Server provides information about the mappings and the vocabularies (ontologies, thesauri) involved in the ARIADNE LOD Cloud.

The current ARIADNE LOD Cloud provided promising results as shown by the Demonstrators.

**Data Mining**

ARIADNE, facilitates a central web portal that provides access to archaeological data from various sources. Parts of these data are being provided as Linked Data. Users can explore these data using traditional methods such as faceted browsing and keyword search, as well as the more-advanced capabilities that come with Linked Data such as semantic queries. A shared characteristic amongst these methods is that they allow users to explore explicit information present in the data. However, since the data is available in structured and explicit form, Machine Learning and Data Mining techniques has been used to identify implicit patterns that exist within this explicit information.

**NLP**

ARIADNE explored NLP with the aim of making text-based resources more discoverable and useful. Partners involved in these activities have specifically focused on the unpublished reports generated by commercial or rescue archaeology, commonly
known as “grey literature”. Aspects of rule-based and machine learning approaches, the use of archaeological thesauri in NLP, and various Information Extraction (IE) methods have been explored. The Named Entity Recognition (NER) techniques were focused on the general archaeological entities of Archaeological Context, Material, Physical Object, Monument, Place, and Temporal (Time Appellation). The methods proved to be capable of extracting CIDOC CRM elements and in some case studies Getty Art and Architecture Thesaurus concepts, in addition to the native vocabularies. General archaeological NLP (GATE) pipelines for English, Dutch and Swedish have been developed. In addition, experimental pipelines were developed for two exploratory thematic case studies on data integration, where the output is expressed as RDF Linked Data via a CRM based data model. An English language pipeline is available for a numismatic case study. English, Dutch and Swedish pipeline are available for a case study of item level data/NLP integration on a loose theme based around archaeological interest in wooden objects and their dating, as expressed in different kinds of datasets and reports. Both case studies have resulted in interactive demonstrators operating over the ARIADNE Linked Data Cloud. All seven pipelines are freely available as open source ARIADNE outcomes.

Potential Impact:
ARIADNE is an EU-funded Integrating Activity aimed at overcoming a situation of dispersed and isolated archaeological data resources. The project has achieved important results that allow data sharing, interoperability and accessibility for research across institutional and national as well as disciplinary boundaries. The evaluation of the achieved results (D2.4) demonstrate that the ARIADNE project has a strong impact in the field of archaeological research and data management in Europe, and that this may become a lasting impact. The results are encouraging and a solid basis for taking appropriate next steps. Based on the project results there is a high potential for further advances based on sharing of data and other resources through the ARIADNE e-infrastructure.

The period of EU support for the ARIADNE project ended. However, the set up of the not-for-profit ARIADNE Association will enable continued operation of the current dataset registry and access portal, further community networking and offering training for potential data providers.

To realise the full potential of the implemented data infrastructure and portal for the archaeological research community, it is necessary to incorporate additional datasets, and ensure the sustainability of the e-infrastructure operation. The platform is a community asset which requires sustained efforts for maintenance and extension. Therefore, all institutional stakeholders in accessible archaeological data, i.e. professional associations, research institutes and funding agencies, should consider using the platform for own and community wide purposes, supporting the platform by making available resources and/or mobilising data providers. Delegation of representatives of the ARIADNE Association should be active to express and promote common interests in the continuation of the e-infrastructure and other activities (i.e. open data advocacy, training and professional development).

The ARIADNE initiative, represented by the ARIADNE Association will focus on providing training for potential providers of new datasets, aligning with the knowledge transfer for new archaeological data archives. Moreover, the Association will maintain the current services by keeping them running, including basic maintenance and updates (if required), until new funding for extension of datasets and services are acquired, for example in a follow-up project. Incorporation of new datasets is also envisaged as a number of institutions have expressed their interest to use the data registry and portal for publishing datasets. This will be achieved on specific arrangements and according to the availability of own or external funds.

Substantial extension of resources and continued sustainability will require a series of funded projects or dedicated national or international funding commitments. An approach for long-term sustainability of the e-infrastructure and services could be a foundation, endowed with renewed significant funding. Another option could be that an ESFRI Research Infrastructure of the humanities and heritage sciences sector (DARIAH, E-RIHS) takes up the e-infrastructure and adapts and utilizes it to support research communities, including archaeological researchers.

Dissemination activities
The ARIADNE consortium consists of partners in sixteen countries including Sweden, United Kingdom, Ireland, Germany, Austria, Hungary, Czech Republic, Slovenia, France, the Netherlands, Italy, Spain, Greece, Cyprus, Romania and Bulgaria. The partners have been very active in disseminating news about the project by:
¥ Giving presentations at national and international events
Transnational Access and training

ARIADNE has offered a range of opportunities for transnational access to the infrastructure and training to archaeologists throughout the project. This has included opportunities for online access to datasets and services, training events, summer schools, group and individual access visits to the physical laboratories of ARIADNE partners.

The feedback that the project has received from participants in these activities has been overwhelmingly positive. Participants highlighted how they valued the opportunities to learn, to network and collaborate with other researchers, and to receive practical help and advice on their research projects.

The transnational access and training activities delivered by ARIADNE fulfilled their objectives of engaging participants with the research infrastructure and delivering on important learning outcomes with regards to the creation, management, access and preservation of archaeological datasets.

During the forty-eight months of the project, ARIADNE has provided online access to archaeological datasets. Initially access was provided to services delivered by its partners (the ADS catalogue, FASTI-online, ARACHNE and Zenon). With the launch of the ARIADNE portal, ARIADNE is now offering integrated online access to datasets from 16 partners in 15 countries.

Monitoring of the services confirms significant research interest in archaeological datasets online. The website statistics for the four external services show consistent visitor traffic throughout the four years of the project, while those for the ARIADNE portal show a strong growth in visitor traffic as the available content has increased and new services have been launched.

ARIADNE dissemination activities have had a visible impact on the number of visitors and page views on the online services. Major conferences, project sessions, workshops and other activities all played their part in contributing to use of these online services. The project partners’ use of social media has also influenced traffic to the individual websites, seen through referrals from Twitter, Facebook and Wikipedia.

Since the launch of the ARIADNE portal on March 2016, traffic to online services provided directly by the project has increased. Visitor traffic to the Fasti Online and ARACHNE services appear to have declined slightly, however traffic to the ADS Portal has increased. ADS datasets were among the first to be ingested to the portal, making up 80% of all the content, and as a result ADS received the largest number of visitors (800) referred from the ARIADNE Portal of the three services.

Comparison of the web-site statistics for the online services reveals that the ARIADNE portal and website, and FASTI online have international audiences with visitors from North and South America, Asia, Australia and Africa as well as Europe. ADS and ARACHNE have strong national audiences but their user bases are also international. Analysis of the browser languages chosen by users also underlines the international nature of the audience for online archaeological data.

User feedback on the ARIADNE Portal has been very positive. Throughout 2016 the content of the portal has been growing and users have been discovering the services that are now available. The audience is growing and, as our survey confirmed, the convenience of being about to use the portal to access to several sources of content is appreciated.

Dissemination Plan

The project dissemination plan described in D4.2 defined the following groups amongst the ARIADNE stakeholder community:

- Internal stakeholders in the partner institutions who have an interest or involvement in archaeological research or management responsibilities relating to project activities;
- Research institutions active in the field as represented by managers and senior researchers with management duties such as
as deans, directors etc.;
- Scholars, researchers and students in archaeological disciplines, field archaeologists and the wider scientific community;
- International networks and research infrastructures in related disciplines;
- Policy makers and policy bodies, and funding agencies including the European Commission;
- Media and the public at large.

Throughout the project partners have worked to raise awareness of the project amongst each of these groups. This has been achieved by updating the website and tweeting regularly, the Newsletters, presenting papers and organizing workshops at national and international conferences, publishing project deliverables, presentations and other materials on SlideShare and other related dissemination activities such as poster sessions, videos on YouTube etc. The numbers of website visitors, newsletter subscribers and Twitter reach, have grown steadily throughout the project. ARIADNE has actively engaged with international networks and research infrastructures from its launch and through out the project. DCH-RP (Digital Cultural Heritage Roadmap for Preservation), DARIAH (Digital Research Infrastructure for the Arts and Humanities), EHRI (European Holocaust Research Infrastructure), CENDARI (Collaborative European Digital Archive Infrastructure) and the European Association of Archaeologists were involved in the ARIADNE’s launch event. DARIAH, CENDARI and CLARIN (Language Studies). The projects regularly exchanged news and supported each other’s dissemination activities.

The ARIADNE dissemination plan has identified six main objectives, together with the corresponding activities.

Objective 1
Description & planned activity
Establishing the project website and portal:
- Designing and building the project website and social media accounts
- Developing the project website adding new content
- Preparing for the launch of the integrated portal and registries

Actual Activity
The website was established in the first month of the project and has been developed and improved throughout. News, project reports, presentations and other content have been added to the site regularly. During 2015-16 the major activities have been the launch of the ARIADNE services and portal, which have greatly increased overall visitor traffic to the sites.

Objective 2
Description & planned activity
Extending the stakeholder database:
- Building the contact database
- Developing the project’s present on the social networks.
- Cooperating with existing communities such as EAA, CAA and research infrastructures

Actual Activity
The project continued to be active in disseminating news, participating in events and establishing collaborations with research infrastructures, institutions, EAA, CAA and others. A bibliography for the project has been established on Zotero. A substantial number of project deliverables and presentations have been uploaded to Slideshare where they have attracted a sizable number of views. Partners have been active in sharing news about the project via their websites, newsletters, and social media account.

Objective 3
Description & planned activity
Informing the stakeholder community about news, events, project activities and transnational access to the infrastructure:
- Posting news and information regularly via the website and social networks.
• Producing the project newsletter
• Disseminating project presentations and documents via SlideShare
• Press notices

Actual Activity
The project has disseminated project news, calls and other information regularly via the website and the social networks.
Social media is driving traffic to specific pages on the website, for example the calls for TNA access. 9 issues of the project newsletter have been produced. Twitter has been used regularly to disseminate project news, calls and other information. The impact of retweeting of ARIADNE news by partners and other followers dramatically extends the reach of the stakeholder database to around 160,000. Social media is driving traffic to specific pages on the website, for example the calls for TNA access. 91 presentations and 27 documents have been uploaded to SlideShare during the period. 3 press notices were prepared to announce the project launch, the Research Infrastructures conference in Rome and the final project event.

Objective 4
Description & planned activity
Informing the research community about transnational access and training opportunities:
• Establishing the transnational access selection panel.
• Preparing the calls to researchers to put forwards proposals for access.
• Use of dissemination channels to advertise training opportunities to researchers.

Actual Activity
External experts were invited to participate in the TNA user selection panel again. The 2014, 2015 and 2016 calls for TNA access were prepared and published on the project website. The calls have been widely disseminated to researchers. Materials and user feedback from the TNA have also been disseminated.

Objective 5
Description & planned activity
Developing the set of dissemination materials:
• Project brochure
• Templates for case studies and fact sheets
• Project poster

Actual Activity
The set of project dissemination materials has been maintained and updated. Flyers have been produced for specific events. Two project booklets have been printed for dissemination at project events, digital copies have been published online. The first project booklet has received more than 6000 views on SlideShare.

Objective 6
Description & planned activity
Presenting the project at relevant international and national events:
• Project presentations
• ARIADNE Workshops and conference sessions
• Final ARIADNE conference

Actual Activity
The project was presented at more than 250 international and national events during the project, reaching more than 12,500 attendees. In addition, around 425 researchers have participated in short training courses and workshops organized by ARIADNE.
The project has exceeded the targets established in the initial dissemination plan:

- More than 135 different institutions have been actively involved in ARIADNE by becoming associates, participating in bilateral meetings, sending researchers to participate in ARIADNE TNA and training events, taking part in user surveys and other activities (the target was 100 institutions).
- At least 13,000 users have participated in events where ARIADNE’s results were presented (the target was 250 individuals). Partners have participated in c. 150 international conferences presenting the project to around 4,500 researchers.
- The project website received 36,611 visitors in 53,849 sessions (the target was 12,000)
- The ARIADNE portal received 10,819 visitors in 15,400 sessions between 1st January and th January 2017 (the target was 800).
- ARIADNE’s social networks have 11,500 members with a reach of around 160,000 followers (the target was 15,000 members). The project newsletter has 410 subscribers (the target was 300).
- The Guides to Good practice received around 2,000 unique page views (the target was 1,500 visitors).

List of Websites:
www.ariadne-infrastructure.eu

Related information

| Result In Brief | Better integration for archaeological data |

Reported by

PIN SOC.CONS. A R.L. - SERVIZI DIDATTICI E SCIENTIFICI PER L UNIVERSITA DI FIRENZE
Italy

Subjects

Scientific Research

Last updated on 2017-05-11
Retrieved on 2019-07-18

© European Union, 2019