ARIADNEplus

D12.1 Initial report on JRA activities

Version 1.0 *(final)*

Date 28 June 2020

Grant Agreement number:	823914
Project acronym:	**ARIADNEplus**
Project title:	**Advanced Research Infrastructure for Archaeological Dataset Networking in Europe - plus**
Funding Scheme:	**H2020-INFRAIA-2018-1**
Project co-ordinator name, Title and Organisation:	**Prof. Franco Niccolucci, PIN Scrl - Polo Universitario "Città di Prato"**
Tel:	**+39 0574 602578**
E-mail:	**franco.niccolucci@pin.unifi.it**
Project website address:	**www.ariadne-infrastructure.eu**

Author: **Carlo Meghini, CNR ISTI**

Contributing partners: **Alessia Bardi, CNR ISTI**

Pasquale Pagano, CNR ISTI

Maria Theodoridou, FORTH

Johan Finn, SND

Document History

- 6.5.2020 - Document skeleton created by Carlo Meghini

- 6 to 20.5.2020 – Contributions by WP leaders of WP12 to WP16

- 21.5.2020 – First draft created by Carlo Meghini

- 10.6.2020 – Final version created by Carlo Meghini and submitted to project management for quality check

- 25.6.2020 – Quality Control check by Holly Wright.

- 28.6.2020 – Final version submitted to EC

Licensing Notice

Table of Contents

List of Figures

1 Executive Summary

This deliverable describes the progress of JRA activities (WP 12-16) carried out during the first year of the ARIADNEplus project, as requested by task 12.6. It offers an integrated overview of all the technological developments of the project in the first 18 months of work, pointing out any delay or relevant issues.

After introducing the objectives of the JRA work packages in ARIADNEplus, and the structure of the workplan (Section 2), Section 3 gives an overview of the progress, focusing on the four main areas of work during the first 18 months of the project: the ARIADNEplus Infrastructure, the data aggregation, the AO-Cat ontology and the development of the Portal. The detailed work performed in each JRA Work Package is detailed in the Sections 4 to 8, each devoted to a single Work Package. These workpackages include:

- WP12 JRA1 - Data Integration and Interoperability (Lead beneficiary CNR)
- WP13 JRA2 - ARIADNEplus Infrastructure Operation and Management (CNR)
- WP14 JRA3 - The ARIADNEplus knowledge management system (FORTH)
- WP15 JRA4 - Innovative Services for Users (SND)
- WP16 JRA5 - Innovative Methods and Pilots (INRAP)

The Sections are structured in a homogeneous way to facilitate clarity. This structure reports on state of each of the above WPs with regard to Goals, Planned Schedules, Results Obtained in the First 18 months, Problems Encountered and Open Issues, and the Outlook for future progress. Each one also includes a table, reporting the status of the activities and the plans for the next period. Almost all the activities are on schedule, with very few exceptions, due to reasons explained in the related section.

The JRA activities of ARIADNEplus are proceeding according to the planned schedules, and at this time the objectives set for the first two years of the project are on track to be fully achieved. This is a major result, in light of the unforeseen necessity to re-design the aggregation process and revise the ontology upon which the ARIADNE Content Cloud is based, but in spite of this, the new ARIADNE Portal has been made available to the members of the Consortium in March 2020, with less than three months delay. A more detailed account of the JRA activities will be given in the five deliverables expected at M24, one for each JRA Work Package.

2 Introduction and Objectives

The present deliverable describes the progress on the joint research activities (JRAs) of the ARIADNEplus project at the end of the first reporting period (Months 1-18). It provides an integrated overview of all the technological developments that have taken place across the Work Packages devoted to JRA. These Work Packages are:

- WP12 JRA1 - Data Integration and Interoperability (Lead beneficiary CNR)
- WP13 JRA2 - ARIADNEplus Infrastructure Operation and Management (CNR)
- WP14 JRA3 - The ARIADNEplus knowledge management system (FORTH)
- WP15 JRA4 - Innovative Services for Users (SND)
- WP16 JRA5 - Innovative Methods and Pilots (INRAP)

The deliverable summarises the current status of each WP, but will be further detailed for the individual deliverables, in the reporting to be produced around the mid-term of the project (Month 25). For convenience, these deliverables are listed in the following Table:

Deliverable No.	Deliverable Title	Type
D12.2	Mid-term report on data WP12 integration - JRA1	Report
D13.2	VREs Operation Mid- term Activity Report - JRA2	Report
D14.1	Mid-term interim report on the ARIADNEplus knowledge management system - JRA3	Report
D15.1	Mid-term interim report on ARIADNEplus services - JRA4	Report
D16.1	Mid-term interim report on ARIADNEplus pilots - JRA5	Report

The present deliverable corresponds to the activities of task 12.6.

Given the objectives of the deliverable, the document has been designed to be a quick reference guide, concise and homogeneously structured. The detailed descriptions of the activities of the involved Work Packages (WPs) will be given in the deliverables reported in the previous Table.

The document consists of one main Section (Section 2) providing an overview of the progress in the JRA WPs, followed by one Section for each one of the WPs from WP12 to WP16. Each WP Section has the same structure, given by:

1. Goal
2. Planned schedule of the activities (table)
3. Results obtained in the first 18 months (bulleted list)
4. Problems encountered and open issues, if any
5. Outlook

Finally, Section 8 concludes the document.

3 Overview of progress

This Section provides an overview of the progress achieved in the first 18 months of the project. It is divided in two sub-sections: the first recapitulates the workplan for the JRA WPs, as it has been presented in the project proposal, while the second identifies the main areas of work and elaborates on the progress achieved in each area.

3.1 Workplan

The following Figure (from page 32 of the ARIADNEplus proposal) schematises the Work Plan of the project and indicates the role of the different JRA WPs in relation to the other WPs.
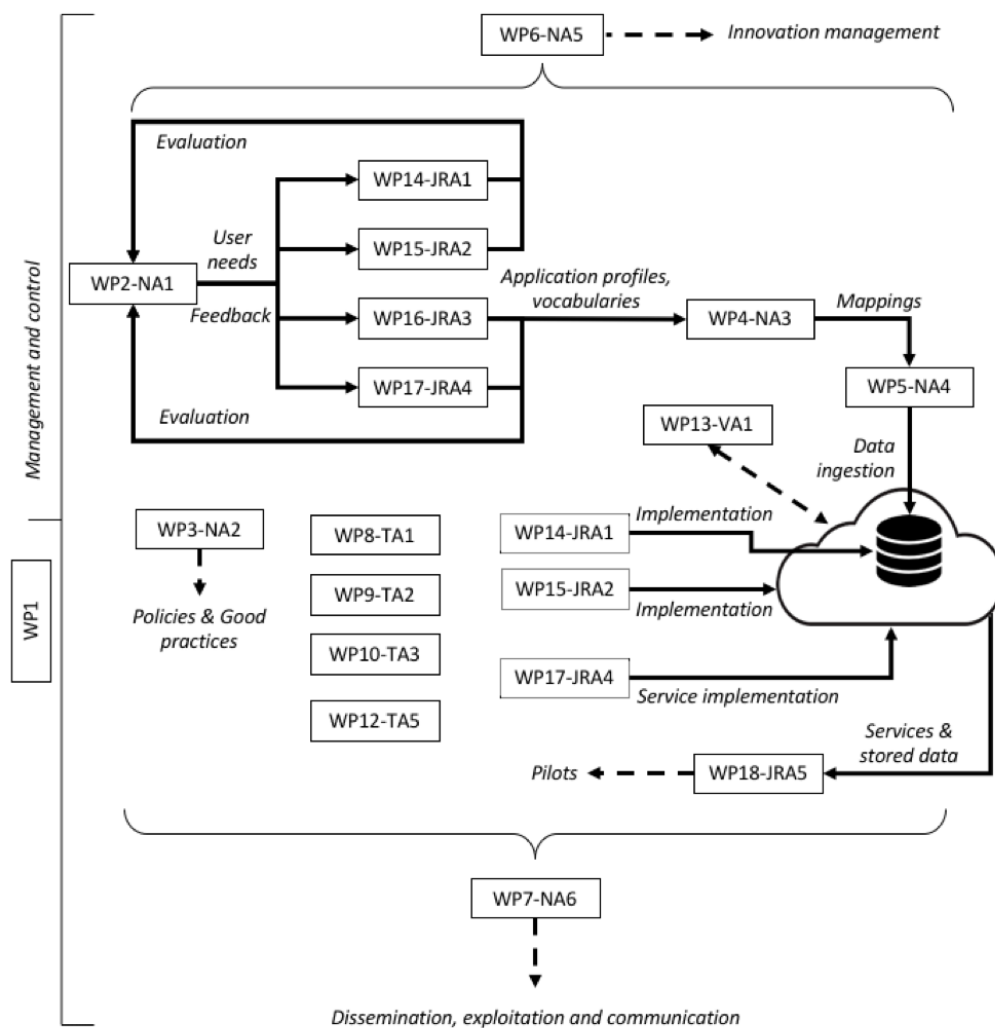

Figure 1. PERT of the ARIADNEplus project.

In brief,

- In WP12 the ARIADNEplus Aggregative Infrastructure, which is called the ARIADNE Content Cloud and forms the main knowledge base of the project, is now set up and operating.

- In WP13 the ARIADNEplus Data Infrastructure was instantiated within D4Science, configured and then further developed to serve the needs of the archaeological community. Moreover, the Virtual Research Environments offering the services to support the R&D activities of the project were created on top of the Data Infrastructure, duly instrumented and put into operation.

- WP14 focuses on the creation of the ARIADNEplus Knowledge Organization System (KOS), defining application profiles (used in WP4 (NA3) for mappings, and generally for managing the KOS), vocabularies and gazetteers.

- WP15 oversees implementation of services, including those which provide useful aids visual, annotation, text mining and geo-temporal tasks and data. It also implements the back-office services that manage queries.

- WP16 implements a number of innovative pilots based on the ARIADNEplus technology and dataset integration.

The timing of the WPs follow a simple rule: All WPs start at Month 1 with preparatory activities (planning, internal WP organisation, design, etc.) and each had an individual WP kick-off meeting at the project Kick-Off. They end at Month 48 contributing to the final reporting, including the project Progress Report.

3.2 Progress achieved

The project Kick-off took place in Prato from February 11 to February 14, 2019. All WPs had a preliminary meeting at the kick-off event of the project as planned. Communication between WPs has been continued at the Steering Committee meetings:

- Steering committee face-to-face Meeting: Bern, September 3, 2019
- Steering committee Virtual meeting: April 14, 2020

The first Joint Technical Meeting was organised to be held in Pisa on March the 2nd and the 3rd. Due to the restrictions caused by the COVID-19, the meeting had to be cancelled and will be re-scheduled. Since the restrictions were set at the very last minute (airport closure announced in the last week of February), it was impossible to move to full virtual meeting, including the hands-on training, at such short notice. However, an online conference was organized on the 2nd of March to partially cover the content.

Progress in the JRA WPs has been carried out on four main fronts:

- Infrastructure and VREs setup and operation
- Data aggregation

- The AO-Cat ontology
- Portal development

A separate sub-section within this document is devoted to each. There has also been activity in other areas, such as service development, but these activities had a secondary role with respect to the above three, in accordance with the Work Plan. Information about them can be found in the individual WP descriptions.

3.2.1   Infrastructure and VREs setup and operation

During the first three months of the project, the infrastructure for ARIADNEplus was designed, configured, and provisioned. The first two VREs, the ARIADNEplus Project VRE and the ARIADNEplus Mappings VRE were created and made available to the members of the Consortium. This allowed the project to achieve Milestone MS12 *First Virtual Research Environments deployed*, scheduled for M20, 18 months early.

The ARIADNEplus Gateway, which is the unique, human-readable access point to the ARIADNE Infrastructure and VREs, had been already created at the kick-off meeting, and could be presented to the participants, who were also instructed on how to access it.

In the subsequent three months, the tools and procedures governing the release of software, methods, and tools for the ARIADNEplus infrastructure were defined and documented in deliverable D13.1. These procedures provided the instructions and rules for the three main patterns governing the provision of software and tools to the ARIADNEplus infrastructure. The first pattern is related to the provision of software to the gCube infrastructure enabling technology. The second pattern is related to the provision of software methods elaborated by the ARIADNEplus community, requiring execution within the ARIADNEplus infrastructure. The third pattern is instead related to the provisions and integration of tools into the ARIADNEplus infrastructure.

By Month 9, the Vocabulary Matching Tool was integrated into the ARIADNEplus infrastructure by hosting it in one of the D4Science sites.

A new VRE for Aggregation Management has been launched on March 19, 2020.

A more detailed account of the Infrastructure, including statistics on the accesses to the VREs, is given in Section 5.

3.2.2   Data aggregation

Data aggregation is essentially an inter-WP activity, carried out by WP12 in cooperation with WP14, and two WPs in the NA set of activities: WP4 Integrating the datasets of the Archaeological Research Communities and WP5 Extending the ARIADNEplus Data Infrastructure (ADI). The objective of this activity was to setup, instrument and operate the ARIADNE Aggregative Infrastructure in order to create the AC.

At the beginning of the project it was decided to re-perform the aggregation of the data already aggregated by the ARIADNE project, due to two main reasons:

1. The enhanced scope and functionality of ARIADNEplus required the development of a more adequate ontology than the ACDM (the ARIADNE Catalogue Data Model), better aligned with the CIDOC CRM and more tightly related to the objectives of the project. As a consequence, the original data had to be transformed into the new ontology, which requires a new aggregation process.
2. The infrastructure for the aggregation of ARIADNEplus metadata is different from that used in ARIADNE, based on different technology and more tightly integrated with the D4Science factory. As a consequence, the inevitable update of the already aggregated data required the re-definition of the aggregation process for these data, which made it effortless to re-aggregate the old data as well.

The operational launch of data aggregation happened at the *First Aggregation Workshop*, held in Pisa on the 13th and the 14th of May, 2019. The workshop was mainly attended by representatives of the four relevant WPs mentioned above and achieved progress on a number of fronts:

- The Aggregation Task Force was launched at the Workshop, and includes representatives of the four WPs involved in data aggregation.
- A programme of work, as well as a series of tools for communication were agreed
- The high-level architecture of the ADI was defined and agreed with project partners.
- Discussion on the aggregation workflows was begun.

The aggregation workflows were finalised in the next few months, to be presented at the *Second Aggregation Workshop*, held in Pisa on the 23rd and the 24th of September 2019, and was attended by a significant number of the content providing partners. The workflows are described in the *Data Aggregation Pipeline User Manual*, an internal document on data aggregation co-authored by representatives of the four WPs involved in data aggregation.

As of today, the project has achieved the aggregation of most of the data provided by ADS, which is by far the largest content provider of the project.

Regular bi-weekly conference calls are held amongst representatives of the four relevant WPs to monitor the progress of data aggregation since the beginning of the COVID-19 pandemic.

3.2.3 The AO-Cat Ontology

The ARIADNE Content Cloud (AC for short) is the information repository collecting all the knowledge managed by the ARIADNE infrastructure. The AC is a composite knowledge base that can be understood as consisting of several related parts, namely:

- the ARIADNE Catalogue, giving general knowledge about the resources of the AC; data resources will be represented in the Catalogue at the collection level;

- several sub-knowledge bases, each containing knowledge statements about a specific archaeological domain.

The ARIADNE Ontology (AO for short) is the ontology providing the terms for making the statements in the AC. The AO is a formal ontology of the resources aggregated by the research infrastructure, with a special focus on the archaeological domain and on its infrastructure, ARIADNE, developed by the ARIADNEplus project.

Mirroring the structure of the AC, the AO will also be structured into sub-ontologies, namely:

- a Catalogue ontology, named AO-Cat, providing terms for the statements in the ARIADNE Catalogue;

- several application profiles, providing terms for the domain-specific parts of the AC.

Work on both the AO-Cat and the application profiles was substantial in the first 18 months of the project. The activity related to application profiles is reported in the Section concerning WP14. The activity related to the AO-Cat is briefly summarised here, as the ontology belongs to WP4, which is not a JRA WP, but the technical development of the ontology required an effort that was produced by the JRA component of the project.

AO-Cat derives from the ARIADNE Catalogue Data Model (ACDM for short), employed in the predecessor ARIADNE project to model archaeological resources, and from the PARTHENOS Entities Model (PEM), employed in the PARTHENOS project, to model the resources managed by a research infrastructure. In its present version, AO-Cat is a contraction of the ACDM driven by the conceptualisation underlying the PEM. In addition, AO-Cat inherits from the PEM its tight relation to the CRM, which is used to represent any aspect of archaeological resources not covered by the ACDM or the PEM. Those aspects not covered by any of the three models mentioned so far, will be dealt with by introducing *ad hoc* terms in AO-Cat.

At present, AO-Cat consists of 20 classes and 59 properties. The taxonomy of the top classes is given in the next Figure:
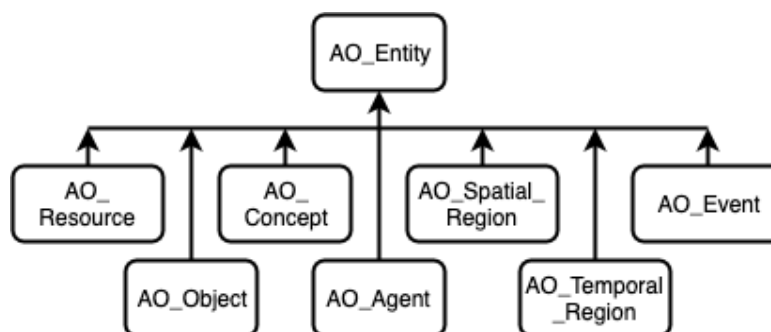

Figure 2. Taxonomy of the top AO-Cat classes.

The AO-Cat is still in development, for this reason it is maintained in two separate versions:

- a version including the basic axioms, expressed in the RDF Schema vocabulary;

- a version including the complete list of axioms, expressed in OWL 2 DL where possible and in first-order logic otherwise.

The first version was used to build the Catalogue and was intentionally kept at a minimal level of specification, so as to avoid generating inconsistencies with the data due to over-axiomatisation. The second version was used to build the ontology proper, and thus includes the complete list of axioms. This duplication implies extra work to keep the two versions aligned, but it avoids creating problems with the aggregation of data. Every time an axiom of the stronger version is invalidated by some data, it is removed from the ontology. Once aggregation is complete, the stronger version will be validated by the data, and released as the official version.

Along with the ontology, we maintain a full mapping to the CIDOC CRM, to insure the widest interoperability of the ARIADNEplus Catalogue with other knowledge bases in the Cultural Heritage domain, and beyond.

The present version of the AO-Cat was included in Deliverable D4.1 *Initial report on dataset WP4 integration* - NA3, delivered at M18. The stronger version of the ontology was included, as it was validated by the data collected so far.

3.2.4 Portal development

The development of the ARIADNEplus Portal is the objective of task T12.3 in WP12. This Task was expected to extend the functionality of the Portal developed by the ARIADNE project, offering search and browse on a larger set of data and adding some (minor) functionality based on the requirements collected from the users in WP2.

During the first six months of the project, the plan of the task was revised to take into account that the old data behind the Portal were going to be re-aggregated according to a different data model, derived from the AO-Cat ontology. This decision, taken by the aggregation team and described previously, required a revision of the technical apparatus of the Portal (data models, APIs and GUI). Furthermore, the version of Elasticsearch employed by the ARIADNE portal had in the meantime become obsolete, and as a consequence, the whole architecture of the Portal required a revision, due to the inclusion of the new version of Elasticsearch.

During the subsequent three months, the design of the new Portal architecture was undertaken, performing an analysis of the fields used by the old ARIADNE portal and their correspondences to the properties of the AO-Cat ontology. In addition, the deployment of a test instance of Elasticsearch 7.4.x was carried out, to start laying the basis of the new architecture. Finally, the adaptation of the portal code to Elasticsearch v7.4.x has been started.

In the next subsequent months, the architecture was finalised, and the adaptation of the code was successfully completed. All source code was updated to PHP7. All queries retrieving the data from the AC were updated to work with the latest version of Elasticsearch. The new Elasticsearch instance was populated with all the content available in the current portal. As a result, the version v1 of the Portal was created.

Portal v1 was then deployed in two instances: "public" and "ghost":

- Public instance replaces the current ARIADNE portal;
- Ghost instance is used by providers and data curators to check their data before going public.

Both instances were live on March 2020, allowing the project to achieve MS6: *New ARIADNEplus Portal launched*. The milestone was achieved with some delay with respect to the expected data (M12) due to the unforeseen plan revision described above.

An important aspect of the Portal implementation consisted of maintaining the search service offered by the old Portal live in parallel, until the new one can replace it. Thus, the public interface to the ARIADNE catalogue will remain the same, until the testing phase of the new Portal is satisfactorily completed. Also due to the recent organisational difficulties caused by the COVID-19 restrictions, and the closure of most academic venues, it is expected that the new Portal will fully replace the old one next summer.

4 Overview on Data Integration

4.1 Goal

The goal of WP12 is to implement and deploy the components of the ARIADNEplus infrastructure that support the integration and interoperability of the data provided by the members of the consortium. The integrated data are made available to the ARIADNEplus portal and the pilots developed in WP16 via two services: (1) the ARIADNEplus knowledge graph, which exposes a SPARQL API, and (2) an Elasticsearch service (see fig. 3.1).
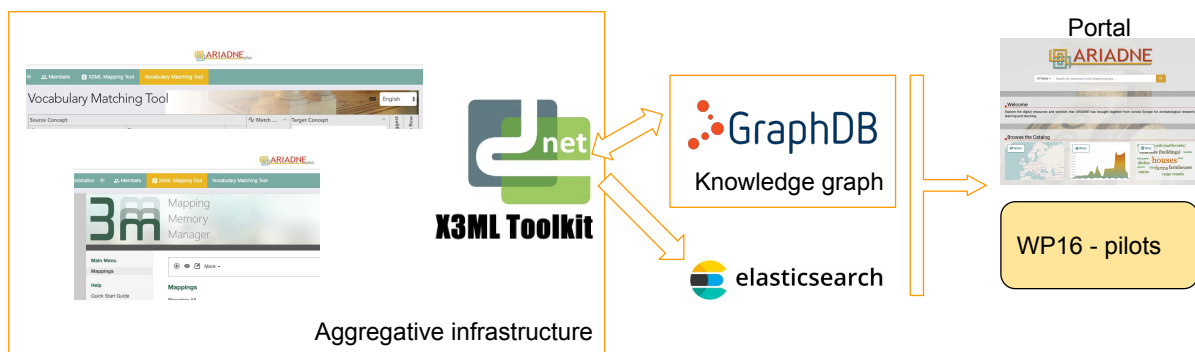
Figure 3. Services and tools for data integration and interoperability in ARIADNEplus.

The aggregative infrastructure includes services and tools required to perform data collection, transformation, and harmonisation, according to domain-specific vocabularies and ontologies:

- 3M Editor: definition of the mappings from local metadata formats to the AO-Cat;
- Vocabulary Matching Tool: definition of mappings from local subject terms to terms of Getty AAT
- D-Net: the ARIADNEplus aggregator is based on the D-Net software toolkit. It collects the provider XML records and integrates the X3ML toolkit for the execution of 3M mappings. It is configured to implement the aggregation workflows defined in collaboration with WP5.

The knowledge graph is a database that supports the Semantic Web paradigm and the Resource Description Framework. It provides a Linked Data endpoint (SPARQL) and reasoning capabilities that can be exploited for the realisation of advanced knowledge discovery services.

The portal is the main entry point for human-actionable search, browse and access the aggregated resources.

4.2 Planned schedule

The following table covers the period M1-M24.

Goal	Planned delivery date	Status (textual description)
3M Editor integrated in AriadnePlus_Mappings VRE	M6 (June 19)	Complete
Vocabulary Matching Tool integrated in AriadnePlus_Mappings VRE	M9 (Sept 19)	Complete
Deployment of the Aggregator (customization of D-Net framework toolkit)	M9 (Sept 19)	Complete. Functionality and customisation subject to WP4 and WP5 requirements (following agile development and continuous delivery paradigms)
Implementation of the aggregation workflow	M10 (Oct 19)	Complete
MS6 New ARIADNEplus portal launched	M12 (Dec 19)	Postponed. First beta version of the new portal available to the consortium in M15 (March 2020). Launch planned for Summer 2020
Availability of the staging and public instances of GraphDB, Elasticsearch and portal to support quality checks	M12 (Dec 19)	Postponed. Staging instances available to the consortium in M15 (March 2020). Public instances will be made available when the aggregated records satisfy the consortium, by M24 (Dec 2020)
ARIADNEplus portal updated with additional features and new framework	M24	Work in progress

4.3 Results obtained in the first 18 months

4.3.1   Aggregative infrastructure

- D-NET framework toolkit instantiated for ARIADNEplus and configured to implement the aggregation workflow devised in WP5. Tests have been run on data provided by ADS.

- 3M Editor updated to allow the generation of RDF records compliant with AO-Cat

- Vocabulary Matching Tool for the generation of mappings from local subjects to Getty AAT deployed on the D4Science infrastructure

4.3.2   Knowledge graph

- Selection and deployment of the service for the knowledge graph (GraphDB, free edition)

4.3.3   Elasticsearch

- Elasticsearch server v7.4.0 available on the D4Science infrastructure and configured by migrating the schema of the old server to the new format

- Data available on the old ARIADNE portal migrated to the new Elasticsearch server

4.3.4   Portal

- Source code of the ARIADNE portal migrated to PHP7

- Queries updated to work with the latest version of Elasticsearch

- Docker setup created for easy deployment

- A user requirements group was formed to formulate the feature requirements and priorities regarding the development of the portal.

4.4 Problems encountered and open issues

The milestone MS6 (Launch of the new ARIADNEplus portal) was postponed due to additional effort not foreseen in the DoA. In particular, the upgrade to the new version of Elasticsearch required more changes to original source code for the portal than anticipated, e.g. upgrade all source code to PHP7. Lack of person months assigned to SND for portal development was resolved by moving resources from WP13 and WP15 to WP12.

The availability of a stable aggregative infrastructure also took longer than expected as the extensive tests conducted on ADS data revealed the need for additional integration and harmonisation tools to be integrated. The evaluation of the performance of the free version of GraphDB is still ongoing, and may lead to the decision to use the commercial version.

4.5 Outlook

4.5.1 Portal

The new version of the ARIADNEplus portal builds upon a new framework called Symfony. The previous portal was built using the Laravel framework. Laravel was chosen due to other requirements in the project, but for current and future purposes, it required significant, unnecessary overheads. We expect the Symfony framework to work better going forward.

A user requirements group was formed to formulate the feature requirements and priorities regarding the development of the portal. This group will analyse the original portal to find areas of improvement, and formulate requirements for new features requests.

The development of the portal will be incremental, and continuous improvements will be made throughout the duration of the project.

4.5.2 Elasticsearch

The Elasticsearch mapping (aka schema) will be updated according to input from the portal user requirements group.

The set-up of the "ghost" and "public" instances are planned to be completed by M24.

4.5.3 Aggregative infrastructure

Inclusion of additional collection protocols for the integration of data coming from sources not compliant with the standard protocols already supported by the aggregator will also be integrated.

4.5.4 Knowledge graph

The performance of GraphDB will be monitored, resources and configuration adapted, and if needed, a plan for the migration to the commercial version of GraphDB will be proposed to the JRA supervisor.

The set-up for the "ghost" and "public" instances are planned to be completed by M24.

5 Overview of Infrastructure and VRE Operation

5.1 Goal

The goal of WP13 is to deliver and operate the ARIADNEplus Infrastructure.

The infrastructure was built by exploiting the computing and stored resources operated and provisioned by D4Science.org together with services for their management and administration. The data, tools and services were deployed on these resources, and made available to the research communities of the project for access and use, via an authentication and authorisation mechanism (also provisioned by this task) compliant with the EOSC identity federation.

The activities performed include:

1. the set-up of monitoring, alerting, and accounting services for all federated resources to guarantee the required Quality of Service (QoS) and
2. the operation of a number of Virtual Research Environments (VREs) providing support for the exploitation of the provisioned storage and computing facilities to the other JRA Work Packages of the project, via the production of training material and how-tos aimed at simplifying all phases of the preparation of new tools and services candidates for integration.

The above-mentioned storage and computing facilities are meant to be accessible via VRE - exploiting the procedures and tools already validated and used in D4Science - by the tools and data required by the services developed in the context of WP15 (ensuring that the proper allocation of resources is guaranteed for their operation) and by WP16 activities.

As a complement, WP13 manages the software release process covering all stages from integration, through documentation and validation, up to provisioning in JRA work packages. Therefore, it defines release and provisioning procedures, establishes the release plan, coordinates the release process, and operates the tools required to support the release and provisioning activities by, also taking care of the distribution of the software.

5.2 Planned schedule

Goal	Planned delivery date (month, e.g. M25)	Status (textual description)
ARIADNEPlus e-infrastructure Gateway	M1	Complete
MS12 First Virtual Research Environments deployed	M20	Complete

5.3 Results obtained in the first 18 months

- Deployment and operation of the ARIADNEPlus e-infrastructure Gateway available at https://ariadne.d4science.org, (from M1);

- Integration/OnBoarding of FORTH's X3ML Mapping Tool service;

- Integration/OnBoarding of the University of South Wales's Vocabulary Matching Tool Service;

- Deployment and operation of the ARIADNEplus_Project VRE, conceived to be the working environment supporting cooperation and collaboration among the ARIADNEplus project beneficiaries (from M1);

- Deployment and operation of the ARIADNEplus_Mappings VRE, conceived to be the working environment supporting the metadata mappings within the project;

- Deployment and operation of the ARIADNEplus_AggregationMgmt

- VRE, where selected project partners discuss data integration issues and procedures to activate or propose to the whole consortium;

- Deployment and operation of the GeoNA Prototype VRE, conceived to be the working environment for the integration, validation, harmonisation, visualisation, and access for archaeological georeferenced datasets collected in Italy. This prototype is intended to be a pilot for any national archaeological geoportal;

- Deployment and operation of the Archeomar VRE, conceived to be the working environment for the visualisation and controlled access of archaeological georeferenced datasets. These datasets contribute also to the GeoNA Prototype VRE;

- Deployment and operation of the environment required for the development and validation of the new version of the Ariadne Portal;

- Deployment and configuration of two accounting dashboards. The first dashboard provides an analytical view of the accesses and usage of the services and VREs; the second provides a geographical analysis of the distribution of users across countries;

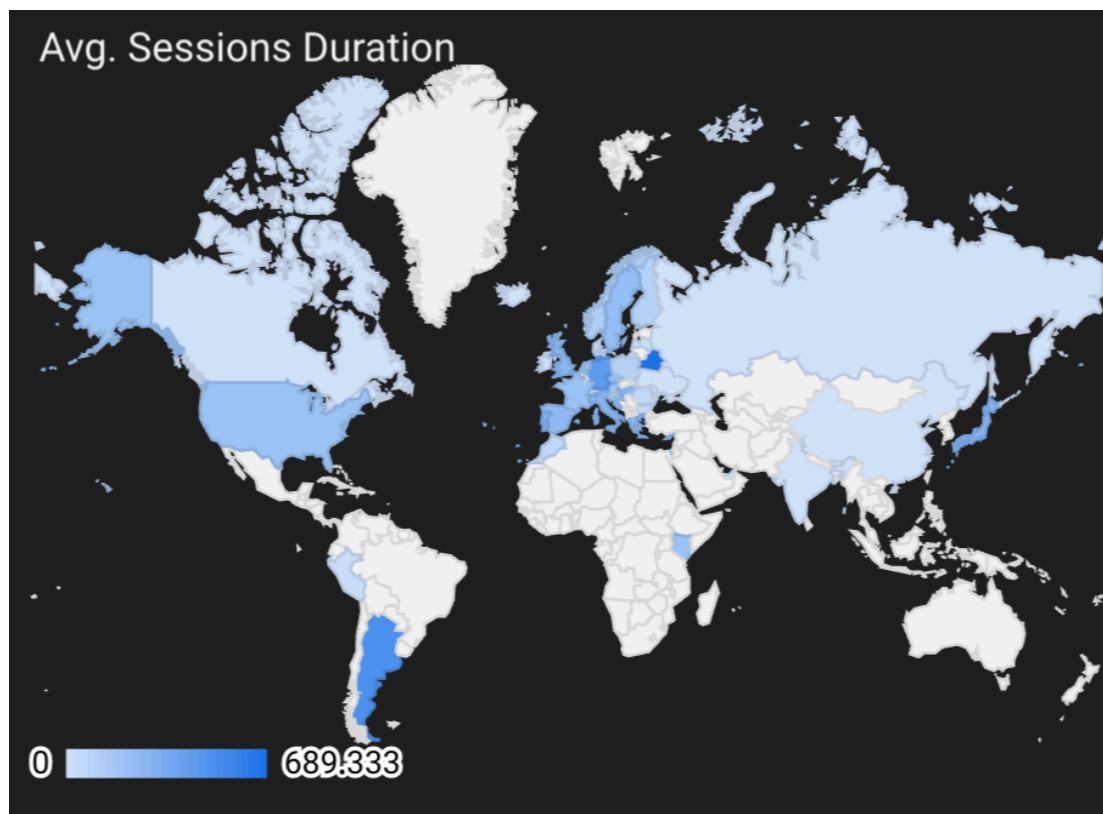- 224 VRE users have registered (as of May 2020) to use the infrastructure, with an average of 350 visits monthly.

Figure 4. Diagram and map of the detailed VRE accesses, and Average Session Duration of the ARIADNEplus Data Infrastructure.

5.4 Problems encountered and open issues

No specific problems or open issues.

5.5 Outlook

Goal	Planned delivery	Status (textual description)
Adoption of the new Identity and Access Management (IAM)	M24	The Identity and Access Management defines and manages the roles and access privileges assigned to VRE users. This activity allows exploitation of the enhanced version of the IAM technology provided by D4Science, and will require the porting of existing users, roles, VRE settings, services privileges, etc. for the technological framework. It will simplify and enable the integration of existing services operated by third-party organisations by reducing the effort required in all technical Work Packages.
Map Viewer	M18	A new version of the map viewer has been designed and implemented to properly visualise and provide access to geographical datasets. It allows an increased and customisable zoom level, tailored Web Feature exploitation, and role-driven access to geographical features. This new version will be deployed and operated in two VREs.
Map Explorer	M25	A new version of the map explorer has been designed. It will allow access to textual descriptions, images, and any other related information associated with a geo package. This new version will be deployed and operated in (at least) two VREs.
Adoption of the new Spatial Data Infrastructure services	M30	The Spatial Data Infrastructure is composed of a set of technologies for the storage, indexing, cataloguing, discovery, and access to geographical datasets according to OGC (Open Geospatial Consortium) standards. This activity will exploit the new version of the SDI provided by D4Science, and will require the porting of existing datasets, styles, VREs settings, access privileges, etc. to the new technological framework.
Adoption of the new Gateway service	M30	A new version of the Gateway service will be designed and rolled out. The new Gateway service will bring enhancements such as the adoption of modern OSGi standards and be broken down into many modules to benefit from the Modular Development Paradigm.

6 Overview of the ARIADNEplus knowledge management system

6.1 Goal

The goal of WP14 is to define the ARIADNEplus semantic framework that:

- monitors the data provision and aggregation process, through a convenient and easy accessed system. This system will offer a flexible and friendly collaborative environment, to be used by many users simultaneously for managing workflow activities and track their progress. The system will act as a reference point for ARIADNEplus stakeholders to watch the status of any process, and provide the appropriate feedback to the respective assignees;

- defines appropriate disciplinary Application Profiles required for the diverse types of data handled by ARIADNEplus. Application Profiles are specialisations expressed as CIDOC CRM extensions for the different thematic groupings, broadly corresponding to those identified in Task 4.4, and aggregated according to similarity;

- provides domain specific vocabularies, a space-time gazetteer and a periodisation system to address the interrelation between space and time along with evaluation of these aspects of the framework.

The validity of the work done in WP14 will be assessed in Task 14.4, in close collaboration with T4.4, and will provide feedback and changes where necessary.

6.2 Planned schedule

Goal	Planned delivery	Status (textual description)
Activity Dash design	M12 (Dec 2019)	Completed
Activity Dash implementation	M18 (June 2020)	Ongoing/mature
Activity Dash testing	M18 (June 2020)	Ongoing. The tool has been demonstrated to WP4, initial feedback received.
Activity Dash Release beta version	M17 (May 2020)	Completed
Activity Dash Release alpha version	M18 (June 2020)	Ongoing/mature
Inscriptions Application Profile initial draft	M13 (January 2020)	Completed
Heritage Science Application profile initial draft	M16 (April 2020)	Completed. It is under test by INFN

Goal	Planned delivery	Status (textual description)
Definition of research protocols for scientific analysis proposed		Ongoing
Investigation and initial development of space-time vocabulary via PeriodO (following T4.4 data integration)		Ongoing

6.3 Results obtained in the first 18 months

6.3.1   Activity Dash

The Activity Dash is a tool to monitor the data provision and aggregation process. The design of the Activity Dash is based on workflows that consist of activities, which in turn consist of tasks. The Activity Dash monitors the data loading process, the mapping process, the thesaurus alignment process, tracking all the available resources and allowing users to collaborate. The Activity Dash offers:

- an easy to access and use environment;

- a convenient way to manage the workflows and activities;

- a collaborative environment for many users to simultaneously add or edit new information regarding processes, and the progress status of activities;

- a means of reliable notification for stakeholders, to be immediately updated when there are any changes applied to workflow activities in which they are involved;

- a reference point for any stakeholder needing to be informed on the progress of any workflow or activity.

6.3.2 Application Profiles

The first year-two sub-disciplines were selected, namely inscriptions and heritage science data. Archaeological partners provided input about the requirements of each sub-discipline, which was synthesised and then verified with domain experts.

Definition of the application profile for inscriptions:

- Assess of available models: CRMtex, EPNet

- Metadata models in use by other ARIADNEplus partners (Inscriptions, graffiti, marks, rock art and other similar material)

- CRMtex: version 1.0 (stable) of the model, submitted for approval to CIDOC CRM SIG

Definition of the application profile for heritage science data:

- Original model developed by PIN in collaboration with INFN

- Support from various partners: CyI, DGPC, LNEC, OEAW

- Specific issues addressed:

  o Sample acquisition and preparation to be properly documented

  o Guarantee of the quality of samples used for analysis

6.3.3 Vocabularies and gazetteers

Preliminary work with some initial investigation of PeriodO and multilingual possibilities using vocabulary mappings and other multilingual resources (AAT, Wikidata, etc) has taken place. Data partners have been supported when entering their period data into PeriodO.

6.4 Problems encountered and open issues

No specific problems or open issues.

6.5 Outlook

Goal	Planned delivery date	Status
Finalisation and implementation of the Activity Dash. Activity Dash Release v1.0	M24 (December 2020)	Ongoing
Integration of the Activity Dash into the Aggregative Infrastructure	M30 (June 2021)	Ongoing
Provision of the tool to users of WP4 in order to monitor their aggregation activities.	M32 (August 2021)	To be started
Testing and Validation of the tool	M36 (December 2021)	To be started
Revision of the tool based on user feedback	M40 (April 2022)	To be started
Definition of the Inscriptions and Heritage Science Application Profiles	M24 (December 2020)	Ongoing

Goal	Planned delivery date	Status
Integration of the Inscriptions Application Profile into the ARIADNE Ontology	M27 (March 2021)	Ongoing
Integration of the Heritage Science Application Profile into the ARIADNE Ontology	M30 (June 2021)	Ongoing
Use of the Inscriptions Application Profile to provide data to the ARIADNE Cloud	M33 (September 2021)	To be started
Use of the Heritage Science Application Profile to provide data to the ARIADNE Cloud	M36 (December 2021)	To be started
Analysis and development of other application profiles	M37 – M48 (January – December 2022)	To be started
Substantial space-time vocabulary via PeriodO (based on T4.4 data integration)	M24 (December 2020)	Ongoing
Supporting data partners as they enter their specific time periods into PeriodO	Now – M48 (June 2021 – December 2022)	Ongoing
Liaise with the Getty Research Institute on requirements for proposal of specialist ARIADNEplus archaeological concepts and multilingual synonyms for AAT	Now – M36 (June 2020 – December 2021)	Ongoing

7 Overview on ARIADNEplus services

The overall goal of Work Package 15 is to provide useful services using ARIADNEplus data to archaeologists. This includes integrating existing services, developing new features into existing services, as well as development of brand-new services. All services should also be integrated into the ARIADNEplus infrastructure by making them available on the D4Science platform.

The work package consists of 10 individual tasks (described below). The services cover a wide range of features useful to researchers, ranging from front-office services like visualisation services, various types of annotations and documentation, to more back-office oriented services like the multilingual query service used by the ARIADNEplus Portal, for querying the underlying aggregating partner resources.

Task 15.1 is responsible for analysing the existing services available in the previous version of the ARIADNE Portal, as well as services already available to partners and others as open source. The task is also responsible for finding and integrating other possible candidates for services in the ARIADNE infrastructure, and to produce a service design for the other services to use.

Tasks 15.2 through 15.5 cover the front-office services and Tasks 15.6 and 15.7 are the back-office services, which are developed under the ARIADNEplus umbrella.

Task 15.2.1 will build upon the VisualMedia EOSCPilot Science Demonstrator. The VisualMedia service, which enables the display of archaeological information in the form of images and 3D models, will be adapted into the ARIADNEplus infrastructure, and made available as an ARIADNEplus service.

Task 15.2.2 will use the pre-existing 3DHOP service at CNR, which links archaeological documentation to the 3D model of an artefact or monument and visualize it accordingly, and adapt it to the ARIADNEplus infrastructure.

Task 15.2.3 will rework the Ephemera service provided by CyI. The service visualises the layers of archaeological excavations in 3D, together with related documentation. The service will also be adapted to the ARIADNEplus infrastructure.

Task 15.3.1 concerns the development of an annotation tool for archaeological reports and other related texts, building upon existing open source tools.

Task 15.3.2 will extend the image annotation tool DAP, developed by CNR and AMZ.

Task 15.4 will integrate and extend a text mining and natural language processing service for the ARIADNEplus service catalogue. The service will build upon the TEXTCROWD EOSCPilot Science Demonstrator service.

Task 15.5 will implement the usual space-time services present in GIS systems and make it available through the ARIADNEplus infrastructure.

Task 15.6 concerns the development of a multilingual query service used by the ARIADNEplus Portal for querying the underlying aggregated partner resources.

Task 15.7 will implement a geoserver for geographical information provided by partners.

7.1 Planned Schedule

Most of the activities shown in the Table below are still to be planned, therefore their delivery data cannot be specified at the moment.

Goal	Planned delivery date	Status
MS17 Design of innovative ARIADNEplus user services defined	M28	Ongoing
Deployment of the VisualMedia service on the ARIADNEplus infrastructure	M28 – M48	To be started
Deployment of the tool for visual organisation of archaeological data on the ARIADNEplus infrastructure	M28 – M48	To be started
Deployment of the Ephemera service on the ARIADNEplus infrastructure	M28 – M48	To be started
Deployment of the archaeological text annotation tool within the ARIADNEplus infrastructure	M28 – M48	To be started
Deployment of the DAP tool within the ARIADNEplus infrastructure	M28 – M48	Tto be started
Deployment of the TEXTCROWD service within the ARIADNEplus infrastructure	M28 – M48	To be started
Deployment of the space-time service within ARIADNEplus infrastructure	M28 – M48	To be started
Deployment of the multilingual query service within the ARIADNEplus infrastructure	M28 – M48	To be started

Goal	Planned delivery date	Status
Deployment of the geoserver service within the ARIADNEplus infrastructure	M28 – M48	To be started

7.2 Results obtained in the first 18 months

- The VisualMedia service authentication module and visualisation interface was updated to be more easily integrated as an ARIADNEplus service.

- The LiDAR visualisation component was designed and released for the VisualMedia service.

- A definition of a common framework for 3D, LiDAR and image annotation expansion was planned within VisualMedia service to meet the needs of the ARIADNEplus partners involved.

- A beta version of a new extension was implemented locally within 3DHOP. The new feature allows archaeologists to create interactive links ("hotspot") from the digital 3D model to related documentation.

- A tutorial, with the list of all the command utilities and relative description available to navigate, interact and perform geometric analysis of the visualized 3D archaeological excavations within the Ephemera tool has been released.

- Short videos explaining each if the command utilities for the Ephemera tool were released.

- Survey for the preparation of user-trials aimed at testing the Ephemera tool, and getting feedback from several groups of specialised users was launched.

- Norwegian textual patterns for time periods and references to date spans for use in multilingual date span matching tool were created.

- Development and Github deployment of ReMatch multilingual date span matching application with interactive web demonstration were completed.

- The Elasticsearch service from the previous ARIADNE project to the latest version was updated.

- The ARIADNEplus Portal was updated to facilitate the new version of Elasticsearch.

- A pilot GeoPortal gateway for the Italian community was deployed within the ARIADNEplus infrastructure.

- SDI (Spatial Data Infrastructure) services for the Italian GeoPortal were configured and Deployed.

7.3 Problems encountered and open issues

Several of the tasks in Work Package 15 are planned for integration into the ARIADNEplus infrastructure, *i.e.* deployment on the D4Science platform. A workshop for partners, to increase the knowledge of D4Science through hands-on sessions, was planned to be held at CNR in Pisa at the beginning of March. Due to the COVID-19 outbreak, this workshop was cancelled. If an in-person workshop remains difficult to organise due to travel restrictions, other possible alternatives will be explored.

7.4 Outlook

The development of the services is making good progress on all fronts. The partners are very keen on continuing development of the existing services and to integrating them into the ARIADNEplus infrastructure. Several new components – to be used to implement new features in the services – have already been developed and released.

8 Overview of the ARIADNEplus pilots

8.1 Goal

The general goals of the WP are to:

- Define the innovative methods enabled by ARIADNEplus services for archaeological research communities.

- Test the services and innovative methods in pilots using real use cases.

- Demonstrate the advantages of using ARIADNEplus to the archaeological user communities.

- Showcase how the ARIADNEplus data and services is building applications for professionals, heritage managers and the public at large.

The general approach consists of:

- Adopting an End-User perspective.

- Being in line with the first objective of the project: that the infrastructure is used, useful and innovative.

- Creating elements that reinforce the archaeological perspective of the infrastructure.

- Developing pilots demonstrating the innovation potential of the project results, methodology and tools.

- Following a Case study approach.

The pilots will include seven case studies involving fourteen partners).

8.2 Planned schedule

The pilots are planned to be operational towards the end of the project (see table below).

Goal	Planned delivery date	Status (textual description)
Establish a common framework for the ARIADNEplus pilots	M18	To be completed in the next 3 months
MS19 First four innovative pilots launched	M35	To be completed in Year 3

8.3 Results obtained in the first 18 months

- Collective and constant update of the archaeological needs specification for the infrastructure

- Input from the archaeological teams into the framework of the services development

- Architecture drafted for the Common pilot framework

As part of task 16.1. "Establishing a common framework for the ARIADNEplus pilots", the task leader (CyI) developed a common evaluation criteria form to describe the methodological impact of the pilot implementation, especially regarding innovation within the archaeological community. The evaluation criteria is described below.

The evaluation criteria form will guarantee the harmonisation of the pilot implementation and integration within the ARIADNEplus infrastructure. The criteria should also consider engagement of the target communities, in order to fully develop the pilots for those communities.

The form should work also as an evaluation, to describe the pilots and the lessons learned from their use. Specifically, the form will be structured as a reporting scheme that will describe both the methodological impact of the pilot implementation, and the innovation created within the archaeological/cultural heritage communities.

Some of the sections and criteria taken into consideration for use within the form are described below. This is a work in progress and will be further implemented and updated throughout the project, according to the different steps of development and final integration of the pilots within the ARIADNEplus infrastructure:

- Description of the pilot (**description**)
  In this section will include a description schema that allows the developers to describe their pilots from a technical point of view (e.g. pilot aims, communities addressed, list and description of command utilities to use the pilot and to interact with the data, tools, etc.).

- Evaluation of the pilot's impact within the archaeological community (**impact**)
  For example, how many users download a pilot (a system to count and evaluate downloads should be considered within the ARIADNEplus infrastructure); how many citations/references a pilot gets, etc. Tracking will be generally useful for all the pilots.

- Evaluation of use (**ease of use** and **ease of access**)
  This will evaluate the ease of access to the pilots and their ease of use.

- Evaluation of the added value of the pilots (**added value**)
  This section will evaluate the added value of the pilots to the archaeological communities, and how the pilots can enhance the functionalities of the ARIADNEplus infrastructure. Specifically, this section will be developed both to evaluate how the pilots can be improved and how they can enhance the functionality of the infrastructure. The added value of a pilot to targeted communities can be assessed both in a general and a specific way. For example, the CyI pilot 'Ephemera' (see subtask 15.2.3 - Visual documentation of an archaeological excavation) can bring general added value to the archaeological community, providing layers of visualisation for a site or structure. For example, a tool like RTI can be a specific added value for the community of the graffiti scholars, in visualising specific written sources on an ancient wall.

- Evaluation of the criteria related to copyrights (**copyrights** and **FAIR principle** fulfilment)
  In line with the overall ARIADNEplus research approach, the criteria form will include a section related to copyright and evaluation of adherence of the pilots to the FAIR principles.

8.4 Problems encountered and open issues

Progress within this work package is dependent on service availability. Partners had few elements with which to work, as details of how the services will be implemented are not yet available.

The draft framework model described above was proposed and commented on at the 2019 Steering Committee meeting, held in Bern (3 September 2019), and it is currently being discussed among involved partners.

8.5 Outlook

The WP will begin in earnest next year. A virtual meeting for the partners involved is planned in the Autumn to refocus the group in time for when services testing needs to begin. It will be an opportunity to review and finalise the common framework within the current draft.

Goal	Planned delivery date	Status (textual description)
Establish a common framework for the ARIADNEplus pilots	M18	A collective seminar (pilot leaders) will be set in 2020 to finalise the common framework of the pilots
MS19 First four innovative pilots launched	M35	Following the above seminar, the pilots will be progressively launched

9 Conclusions

In conclusion, the JRA activities of ARIADNEplus are proceeding according to the planned schedule, and currently it is reasonable to believe that the objectives set for the first two years of the project will be fully achieved.

This is a major result, in light of the unforeseen necessity of re-designing the aggregation process and revising the ontology at the core of the ARIADNE Content Cloud (as illustrated in Section 3 of the present document). This necessity has only impacted the achievement of Milestone MS6, as it was set for M12 and was thus too early to give the Consortium time to absorb the necessary changes. In spite of this, the new ARIADNE Portal was available to the members of the Consortium by March 2020, with less than three months delay.

The impact of the COVID-19 lockdown has been relevant only for meetings concerning specific tasks, and for the Joint Technical Meetings, which included hands-on sessions. Whenever possible, the meetings have been re-organised online, and meetings in person will be re-scheduled when travel is again possible.

A more detailed account of the JRA activities will be given in the five deliverables expected at M24, one for each JRA Work Package.