



GRANT AGREEMENT NUMBER:	693548
PROJECT ACRONYM:	ArchAIDE
PROJECT TITLE:	Archaeological Automatic Interpretation and Documentation of cEramics
FUNDING SCHEME:	H2020-REFLECTIVE-6-2015
PROJECT COORDINATOR	Prof Maria Letizia Gualandi, UNIFI
TEL:	+39 05022 15817
E-MAIL:	maria.letizia.gualandi@unifi.it

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N.693548

Progress Report 1

D1.4

version: 1.3

Revision: Final

Work Package	1
Lead Author (Org)	Gabriele Gattiglia (UNIFI)
Contributing Author(s) (Org)	Julian Richards (UoY), Holly Wright (UoY), Francesca Anichini (UNIFI), Michael Remmy (UCO), Eva Miguel (UB), Roberto Scopigno (CNR), Matteo Dellepiane (CNR), Massimo Zalocco (INERA), Lior Wolf (TAU), Diego Lucendo Diaz (BARAKA), Llorenç Vila (ELEMENTS).
Due Date	31st May 2017 (M12)
Date	12th June 2017



Project co-funded by the European Commission within the ICT Policy Support Programme		
Dissemination Level		
P	Public	X
C	Confidential, only for members of the consortium and the Commission Services	

Revision History

Revision	Date	Author	Description
0.1	15.05.2017	Gabriele Gattiglia	Initial outline, structure of the document
0.2	17.05.2017	Julian Richards, Holly Wright	Draft contents added
0.3	18.05.2017	Matteo Dellepiane, Massimo Zallocco	Draft contents added
0.4	20.05.2017	Michael Remmy	Draft contents added
0.5	22.05.2017	Eva Miguel	Draft contents added
0.6	26.05.2017	Diego Lucendo Diaz, Llorenç Vila	Draft contents added
0.7	01.06.2017	Gabriele Gattiglia, Julian Richards	Draft contents added
0.8	05.06.2017	Gabriele, Gattiglia, Eva Miguel	Draft contents added
0.9	06.06.2017	Lior Wolf	Draft contents added
1.0	08.06.2017	Francesca Anichini, Julian Richards, Roberto Scopigno, Matteo Dellepiane	Draft contents added
1.1	09.06.2017	Gabriele Gattiglia, Massimo Zallocco, Francesca Anichini, Roberto Scopigno, Matteo Dellepiane	Draft contents added
1.2	12.06.2017	Michael Remmy, Francesca Anichini, Llorenç Vila	Draft contents added
1.3	12.06.2017	Gabriele Gattiglia	General Review

Disclaimer

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.



Table of Contents

Abbreviations	4
Executive summary.....	5
1 Explanation of the work carried out by the beneficiaries and Overview of the progress	7
1.1 Objectives	12
1.2 Explanation of the work carried per WP	15
1.2.1 Work Package 1 – Management	15
1.2.2 Work Package 2 - Methodologies, specification and design	17
1.2.3 Work Package 3 - Creation of the Application Database	20
1.2.4 Work Package 4 - Technologies for the digitization of catalogues.....	21
1.2.5 Work Package 5 – Population of the Database	23
1.2.6 Work Package 6 - Shape and image-based similarity search and retrieval	26
1.2.7 Work Package 7 - The mobile tool and Front-end Desktop Application	28
1.2.8 Work Package 8 - Test and assessment of the overall system on application scenarios	28
1.2.9 Work Package 9 - Communication/Public engagement/Innovation	28
1.2.10 Work Package 10 – Archiving and Dissemination	34
1.3 Impact.....	36
2 Update of the plan for exploitation and dissemination of result	37
3 Deviations from Annex 1 and Annex 2	40
3.1 Tasks	40
3.2 Use of resources (not applicable for MSCA)	42
3.2.1 Description of use of resource for direct personnel costs declared as unit costs.....	42
3.2.2 Explanation of the total funding amount used for dissemination and communication activities ..	42
Summary	Errore. Il segnalibro non è definito.



Abbreviations

WP: Work package

M: Month

UNIPi: Università di Pisa

UoY: University of York

UB: Universitat de Barcelona

UCO: Universitaet zu Koeln

TAU: Tel Aviv University

CNR: Consiglio Nazionale delle Ricerche

INERA: Inera srl

BARAKA: Baraka Archeologos S.L.

Elements: Elements centro de gestio i difusio de patrimoni cultural



Executive summary

This document includes an overview of the project results towards the objective of the action in line with the structure of the Annex 1 to the Grant Agreement during the first reporting period from 1st June 2016 to 31st May 2017, including a summary of exploitable results and an explanation about how they can/will be exploited.

The objectives of ArchAIDE are to support the work of archaeologists with innovative computer-based tools, able to provide the user with features for the semi-automatic description and matching of potsherds over the huge existing ceramic catalogues. Indeed, pottery classification is of fundamental importance for the comprehension and dating of the archaeological contexts, and for understanding production, trade flows and social interactions, but it requires complex skills and it is a very time consuming activity, both for researchers and professionals.

These objectives will be achieved through the development of:

- an as-automatic-as-possible procedure to transform the paper catalogues in a digital description, to be used as a data pool for search and retrieval process;
- a tool that will support archaeologists in recognising potsherds during excavation and post excavation analysis, through an easy-to-use interface and efficient algorithms for search and retrieval of the visual/geometrical correspondences;
- an automatic procedure to derive a complete potsherd's identity card by transforming the data collected into a formatted electronic document, printable or visual;
- a web-based real-time data visualization to improve access to archaeological heritage and generate new understanding;
- an open archive to allow the archival and re-use of archaeological data, transforming them into common heritage and permitting economic sustainability.

Currently, the project has ended its first year. Work Packages 1, 3, 4, 5, 6, 7, 9, and 10 are active. WP2 is ended on M8, whereas WP8 (Test and assessment of the overall system on application scenarios) will begin on M22. The project began on 1st June 2016 with WPs 1 Management, 9 Communication/Public engagement/Innovation, 10 Archiving and Dissemination, and 2 Methodologies, specification and design. WP2 was dedicated to define the working scenario in archaeological investigation and pottery processing, the database structure, the selection of the pottery catalogues and of test bed cases. The achievement of MS1 Preliminary report about the archaeological and technical specification of the system on M4 gave the start to WPs 3, 4 and 6. Using the ADS Roman Amphorae database as a point of departure, WP3 Creation of the Application Database has completed the database design and has begun on incorporating multilingual vocabularies of pottery characteristics. WP4 Technologies for digitisation of catalogues evaluated different solutions for OCR, in order to analyse flexibility, amount of languages supported, extendibility. The advanced OCR implementation (that aims at automatically filling the database structure) is currently under finalisation. An automatic application that is able to analyse a scanned version of the drawing has been implemented. In addition, a 3D representation of the drawing is created automatically. The achievement of Milestones 2 First Version of the database ready for being populated and 3 First validated version of the technologies for the digitization of paper catalogues ready to be used for starting DB population on M8 permitted to start WP5 Population of the Database. Once defined the workflow for populating the database, WP 5 began to digitise catalogues and photograph potsherds for training the algorithms. After a first manual population of database to see functionality and bugs,



the database was automatically populated with the data from the ADS Roman Amphorae digital resource. WP6 Shape and image-based similarity search and retrieval has focused on building synthetic training data of 3D vessels as well as simulating breaking them and obtaining matching sherds. This data will be used in order to develop the algorithms necessary for matching sherds with catalogue-based drawings of vessels. While the database of ceramics was populated, an initial overview of the data was done to identify the sort of similarity to be detected. When doing so, it was discovered that there are some problems with appearance based similarity. Consequently, the maximum effort was allocated on shape-based similarity, which allows much more freedom in gaining the data. WP7 The mobile tool and Front-end Desktop Application achieved the Reference Database API for the retrieval of the ceramic types from the mobile tool through the search engine produced in WP6, started the implementation of the core components of the mobile tool, and of the “Result database” (server database supporting the mobile tool). WPs 1, 9, and 10 will last during all the project lifetime. WP1 guarantees the correct implementation of the overall management, the monitoring of the project and its activities, and the achievement of the project objectives. WP9 defined the communication strategies in order to engage a large audience supported by a strong social media communication. The logo and the visual identity of the project have been designed, as well as a line of products to promote the project to a wider public. WP10 started the dissemination and promotion of the results of the project, and worked to make the project collaborative and transparent to researchers in the archaeological domain and beyond.

The project expects impacts on archaeological professional labour market, archaeological research, and training and education in archaeology.

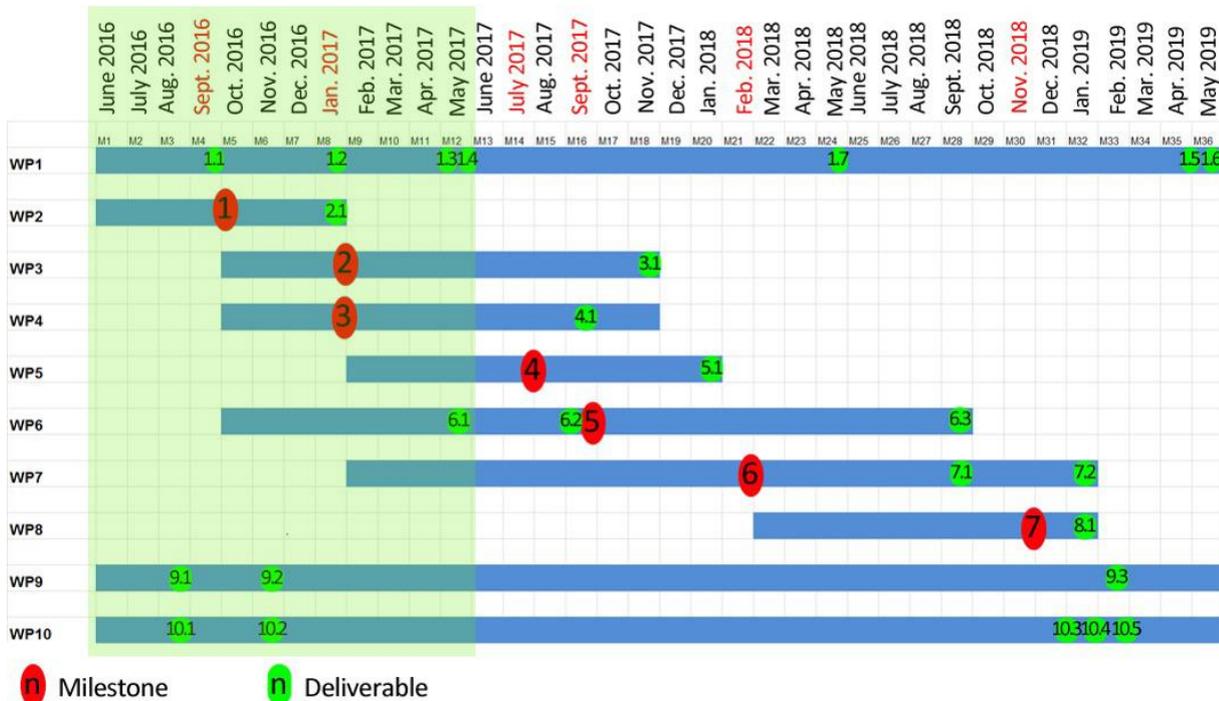
As for the archaeological professional labour market, it is estimated that approximately 33,000 archaeologists now work across Europe, 11,350 in the USA, 6,255 in Japan, and between 500 and 600 in Australia. This information underlines the world-wide potential market of ArchAIDE. ArchAIDE will give to private companies the possibility to reduce time and costs, or to redistribute and optimize the activity of professionals, bringing positive effect on the whole sector.

As for archaeological research, ArchAIDE will move archaeologists from spending time on routine tasks like drawing and classification to create background knowledge and enable new knowledge generation, permitting to develop new research field in archaeology. ArchAIDE allows the creation of new archaeological data that will be made available both through the interactive app (supporting data access and visualization) and published as open data. In this way, ArchAIDE will help to broaden the horizons of archaeological research (new typological studies about archaeological pottery, trade flows and their economic impact, geographical diffusion, etc).

As for education and training, students complete their studies with training activities led by Universities or Archaeological companies. In such a context, ArchAIDE will become an extra training tool.

Besides, we have to consider also the millions of global citizens that approach archaeology in a non-professional manner and in all those related areas (educational, touristic, entrepreneurial, etc) that draw on open archaeological data to produce derivatives content such as storytelling, educational tools, etc. Supporting ordinary people will increase awareness and cultural interest in the population.

1 Explanation of the work carried out by the beneficiaries and Overview of the progress



Currently, the project has ended its first year. Work Packages 1, 3, 4, 5, 6, 7, 9, and 10 are currently active. WP 2 is ended on M8, whereas WP8 *Test and assessment of the overall system on application scenarios* will begin on M22.

The project began with WPs 1 *Management*, 9 *Communication/Public engagement/Innovation* and 10 *Archiving and Dissemination*, which will be discussed later, and WP2 *Methodologies, specification and design*. WP2 was dedicated to define the working scenario in archaeological investigation and pottery processing with a detailed exposition of the main procedures and objectives in the practice of professional archaeology concerning the study of archaeological ceramics, the database structure, the issues related to the Data Preservation, and the selection of the pottery catalogues and of test bed cases. A careful analysis of state-of-the-art provided precious input regarding the main features that could be extracted from the drawings, and used for the description of the typologies and sherds. An in-depth discussion about the functional specification of the platform devices brought to realisation and implementation of a Mock-up of the ArchAIDE application, that has been useful for the beginning of the related WPs. The achievement of the *Preliminary report about the archaeological and technical specification of the system* (MS1) on M4 gave the start to WPs 3, 4 and 6. WP2 ended on M8 with the submission of *D2.1 Detailed system specification, including the representation structure for shape data*.

WP3 is in its 8th month of work and UoY has completed the database design in partnership with INERA. CNR, UCO. UoY coordinated with INERA as the design moved into the implementation phase. Upon completion of the database design, work began on incorporating multilingual vocabularies. Scoping by UoY of the potential for using the Getty Art and Architecture thesaurus (AAT) as a neutral spine for creating multi-lingual mappings of pottery characteristics was completed, with positive results. Mapping of multilingual vocabularies of pottery characteristics to the AAT were completed by UoY, UB, UNIPI and UCO. Mappings were transformed to XSLT by UoY and delivered to INERA in JSON format for implementation. INERA has completed

implementation of the comparative database, based on the design completed by ADS, and partners are carrying out testing. Population of the database with comparative data is underway, and the multi-lingual vocabulary functionality has been implemented by INERA.

Also, WP4 *Technologies for digitisation of catalogues* is in its 8th month of work. Different solutions for OCR have been evaluated, in order to analyse flexibility, amount of languages supported, extendibility. Two different implementation of the OCR-based tool have been proposed. The advanced OCR implementation (that aims at automatically filling the database structure describing by analysing the structure of the description) is currently under finalisation. An automatic application that is able to analyse a scanned version of the drawing has been implemented. In addition, a 3D representation of the drawing is created automatically. The 3D generation tools are currently under use for the initial population of the ArchAIDE database.

The Achievement of Milestones 2 *First Version of the database ready for being populated* and 3 *First validated version of the technologies for the digitization of paper catalogues ready to be used for starting DB population* on M8 as scheduled, and respectively related to WP 3 and 4, permitted to start WP5.

At present, WP5 *Population of the Database* is in its 4th month of work. The main goal of this work package is to populate the database using the technologies developed in WP4. Upon agreement on pottery classes (Amphorae, *Terra Sigillata*, Medieval and Postmedieval pottery) for feeding the database, started the digitization of catalogues, and the taking of picture of pottery sherds for training the algorithms to be developed in WP6. Contemporaneously, the research related to text mining and information extraction and the discussion for defining a workflow for populating the database began. The first manual population of database to see functionality and bugs, was followed by the automatic import of data from the ADS Roman Amphorae digital resource.

WP6 *Shape and image-based similarity search and retrieval* is in its 8th month of work. The task specific work for WP6 has focused so far on building synthetic training data of 3D vessels as well as simulating breaking them and obtaining matching sherds. This data will be used in order to develop the algorithms necessary for matching sherds with catalogue-based drawings of vessels. While the database of ceramics was populated (WP5), an initial overview of the data was done to identify the sort of similarity to be detected as part of this task. When doing so, it was discovered that there are some problems with this data. Deliverable D6.1 was not achieved at the due date, and has been postponed, whereas maximum efforts were allocated to D6.2. These problems were identified and solutions suggested were to shift the focus toward engraved signatures and other appearance based clues, that can be accommodated very efficiently. Shape-based similarity poses similar challenges, but the reliance on shape only and not on texture and appearance, allows much more freedom in gaining the data. Therefore, most of the work was indeed acquiring data, by synthetically generating data. There's still some work on completing the implementation of the alignment algorithm, and after that we expect starting to generate the synthetic sherds to have data for training the machine learning algorithms.

WP7 *The mobile tool and Front-end Desktop Application* is in its 4th month of work, and currently, only Tasks 7.1 and 7.3 are active. The goal of this WP is provide a usable tool following the requirements coming from WP2, integrating the search and retrieval mechanism defined and developed in WP6. During this period, the analysis of the potential scenarios where the mobile tool would be helpful was completed, the technical architecture for the mobile tool and desktop tool defined, and the definition of a detailed mock-up of the mobile tool almost completed. Contemporaneously, the implementation of the core components of the mobile tool started, the Reference Database API for the retrieval of the ceramic types from the mobile tool through the search engine produced in WP6 was achieved and started the implementation of the "Result database" (server database supporting the mobile tool).

Work Package 1, 9, and 10 started at the beginning of the project and will last during all the project lifetime.

WP1 is in its 12th month of work, and all the tasks are in progress as planned. The main objective of WP1 is to guarantee the correct implementation of the overall management, the monitoring of the project and its activities, and the achievement of the project objectives. The first steps were the definition of the Management Structure and the transfer of the first tranche of the respective shares of the financial support received from the EC to all the partners. The Management took care of the organisation of the kick-off meeting as well as the other General and technical meeting. The submission of D1.1 *Quality Plan* (M4) defined the procedures to be used, the decision making protocols, the reporting mechanisms, the organization of meetings, the control of information flow, the reliability of the output, and the preparation of documentation for submission to the EC, the success criteria for each deliverable, and also defines configuration management procedures. The submission of D1.2 *IPR and Knowledge Management Plan* (M8) defined how the consortium is going to deal with Intellectual Property and Access Rights on its foreground and background knowledge, and has been used as a basis for the first draft of the Exploitation Plan. Deliverables D1.3 *Financial Report 1* and D1.4 *Progress Report 1* were realised for the closure of the first reporting period.

At present, WP9 is in its 12th month of work, and all the Tasks are in progress as planned. The main objective of this WP is communicate the implementation of the project, in order to maximise its impact beyond the project community. A *Communication Plan* was realised during M2 and M3 (D9.1), and defined aims and strategies about the communication activities that will be carried out during whole the project, and the choice of using the "visual storytelling" concept to engage a large audience. The Plan has been written simultaneously with the Dissemination Plan (D10.1) and it is closely related to it. The logo of the project has been designed (M1-3), and the visual identity of the project around the logo has been created. A line of products has been designed to promote the project and its aims to a wider public. An internal web site has been defined at the beginning of M3 whereas the public web site was on-line since the end of M6 (D9.2 *Website and promotional kit Plan*). Since M4 a social media communication started and a calendar of publication on the social networks was planned.

At present, WP10 is in its 12th month of work. The objective of this WP is to archive, disseminate and promote the results of the project, and to make the project collaborative and transparent to researchers in the archaeological domain and beyond. D10.1 *Dissemination Plan* was submitted on M3. D10.2 *Data management plan* was completed on M6, in the form of an online document written using the templates within the 'DMPonline' tool: part of the the Open Research Data Pilot (ORD) funded under Horizon 2020. Dissemination activities were carried on at CAA conference in Atlanta, USA, March 14-16, 2017; AIUCD Conference in Rome, 26-28 January, 2017. Next activities are foreseen at EAA conference in Maastricht, 30 August to 3 September, at EMAC Conference in Bordeaux, 6-9 September, 2017, at ICDAR in Kyoto 9-15 September, 2017. Finally, we have started the drafting of D10.5 *Exploitation Plan* (which is synthetically presented in this report and available on request; to be delivered on M33).

Table 1: List of Deliverables. In green Deliverable achieved at the due date. In Yellow Deliverable delayed.

N.	Title	WP n.	Leader	Type	Dissemination level	Due Date
D9.1	Communication plan	WP9	UNIFI	Report	Confidential, only for members of the consortium (including the Commission Services)	(3) AUG.2016
D10.1	Dissemination plan	WP10	UoY	Report	Public	(3) AUG.2016
D1.1	Quality plan	WP1	UNIFI	Report	Confidential,	(4) SEPT.2016

					only for members of the consortium (including the Commission Services)	
D9.2	Website and Promotional kit	WP9	UNIPI	Websites, patents filling, etc.	Public	(6) NOV.2016
D10.2	Data Management Plan	WP10	UoY	Report	Public	(6) NOV.2016
D1.2	IPR and knowledge Management Plan	WP1	UNIPI	Report	Confidential, only for members of the consortium (including the Commission Services)	(8) JAN.2017
D2.1	Detailed system specification, including the representation structure for shape data	WP2	UNIPI	Report	Confidential, only for members of the consortium (including the Commission Services)	(8) JAN. 2017
D1.3	Financial report 1	WP1	UNIPI	Report	Confidential, only for members of the consortium (including the Commission Services)	(12) MAY 2017
D1.4	Progress report 1	WP1	UNIPI	Report	Public	(12) MAY 2017
D6.1	Algorithms for pairwise similarity between ceramics based on appearance (software)	WP6	TAU	Other	Confidential, only for members of the consortium (including the Commission Services)	(12) MAY 2017
D4.1	Final version of the digitization system	WP4	CNR	Demonstration	Confidential, only for members of the consortium (including the Commission	(16) SEPT.2017

					Services)	
D6.2	Algorithms for pairwise similarity between ceramics based on shape and appearance (software)	WP6	TAU	Other	Confidential, only for members of the consortium (including the Commission Services)	(16) SEPT.2017
D3.1	Final release of the database implementation (software)	WP3	UoY	Other	Confidential, only for members of the consortium (including the Commission Services)	(18) NOV.2017
D5.1	Mapping tool	WP5	UCO	Other	Confidential, only for members of the consortium (including the Commission Services)	(20) JAN.2018
D1.7	Progress report 2	WP1	UNIPI	Report	Public	(24) MAY 2018
D6.3	Unified similarity kernel, including tools for results refinement and visualization (software)	WP6	TAU	Other	Confidential, only for members of the consortium (including the Commission Services)	(28) SEPT.2018
D7.1	Data analysis and visualisation component (software)	WP7	INERA srl	Other	Confidential, only for members of the consortium (including the Commission Services)	(28) SEPT.2018
D7.2	Release of the final version of the mobile app (software) and the desktop front-end (software)	WP7	INERA srl	Other	Public	(32) JAN.2019
D8.1	Report of final recommendation for WP8	WP8	UB	Report	Public	(32) JAN.2019

D10.3	Dissemination kit (final release)	WP10	UoY	Websites, patents filling, etc.	Public	(32) JAN.2019
D10.4	Web APIs for REST services	WP10	UoY	Other	Public	(32) JAN.2019
D9.3	Promotional video	WP9	UNIPi	Websites, patents filling, etc.	Public	(33) FEB.2019
D10.5	Exploitation Plan	WP10	UoY	Report	Confidential, only for members of the consortium (including the Commission Services)	(33) FEB.2019
D1.5	Final financial statement	WP1	UNIPi	Report	Confidential, only for members of the consortium (including the Commission Services)	(36) MAY 2019
D1.6	Final project report	WP1	UNIPi	Report	Public	(36) MAY 2019

1.1 Objectives

List the specific objectives for the project as described in section 1.1 of the DoA and described the work carried out during the reporting period towards the achievement of each listed objective. Provide clear and measurable details.

Objective	Completion Criteria	Work carried out during the reporting period
1. Enable scalable and cost-effective documentation of archaeological findings.	The demonstration of a hardware and software system prototype capable to quickly and semi-automatically acquire shape characteristics of archaeological findings (focusing on small scale artefacts and based on digital photography). Documentation will include text and visual representations and the production of drawings. The	The first year of the project was spent in eliciting the requirements of the archaeologists in terms of which information need to be stored, the level of details and the format. Furthermore, a discussion was carried out about the way in which the mobile and desktop tools would integrate in the archaeological processes (excavation, classification, dating, etc.). The result of the activities was the final design of the "Results Database" (to store findings data) in WP3 and the production of preliminary design of the mobile app functionalities (WP7) based on the User Centred Design methodology.

	system front-end will run on a mobile device, to ensure easy use on-site.	M21(early release) M32 (final release)
2. Enable digitization of current catalogues (by automatic processes), including the conversion from raster to digital conceptual models of the artefact classes defined in the catalogues.	The demonstration of a system prototype for the digitization of the catalogue and the production of conceptual models of the classes defined. This is a data acquisition campaign, which has to be completed in an early phase of the project (but requires some algorithmic design and implementation).	<p>The first year of the project has been devoted to set the foundations for the digitization of the paper catalogues, both for the creation of the database and for the use of extracted data for appearance and shape-based classification. The first part of the reporting period was dedicated to an in-depth discussion by the whole consortium in order to:</p> <ul style="list-style-type: none"> • choose the pottery classes and the associated paper catalogues; • select all the basic fields of the database, trying to map the structure of the paper catalogues; • analyse the work by archaeologists to define the geometric and appearance features that can be extracted from the catalogues and from the images taken on-site; • based on the selected catalogues, two main actions towards digitization have been brought on: • for the digitization of text, OCR tools were developed to support the work of WP5; • for shape-based classification, an automatic digitization pipeline, to extract the geometric features and build a 3D representation, was developed and used to populate the database. <p>A first prototype of the tools has been released in M8.</p> <p>M16 (final release) M18 (evaluation)</p>
3. Enable on-the-field support to the archaeologist in the interpretation phase.	The demonstration of an enhanced version of the system running on mobile platforms, which will include also the component supporting the interpretation phase. Instruments for supporting the professional in the visual analysis of the finding, proposing possible matches (based on a search for shape-similarity run over a remote server storing the	<p>The first year of the project has been dedicated to the realisation of the ArchAIDE system Mock-up and to the definition of the technical architecture for the mobile tool and desktop tool. Currently, the implementation of the core components of the mobile tool (communication, user authentication, 2D image acquisition, etc.) started. The work related to data analysis and data visualisation will begin in the next months.</p> <p>M21 (early release) M32 (final release) M32 (evaluation)</p>

	digital catalogue of classes, produced in Obj. 2).	
4. Enable the archival of all data produced on a remote archive, supporting also access through advanced presentation tools.	<p>The demonstration of the complete version of the system running on mobile platforms, which will include the final component for data archival on remote repositories. The system will enable the archival of all data produced: textual data, raw images, models produced (2D drawings, 3D models), results of the interpretation phase, etc. These data will be accessible also via a standard web-based interface; it will provide also advanced visual presentation tools to enable the inspection of the media files.</p>	<p>The first step to achieve the objective was the definition of the “Reference database” where to store the ceramic types to be used for matching and interpretation of excavation findings. This has been done, released, and was started the population by the archaeological partners. The produced system allows the archival of information, depictions (2D, 3D, SVG, etc.) and geographical info about origin, fabrics and occurrences (where was found).</p> <p>Within WP7 was started the analysis and design for the visual presentation of the ceramic types and produced a first prototype of user interface.</p> <p>M21 (early release) M32 (final release) M32 (evaluation)</p>
5. Enable reuse of all data collected and preserved in the open data repository	<p>The realisation of an open data archival policy, which is able to provide long-term preservation and as wide as possible data reuse. The release of web APIs for REST services, in order to allow the creation of new applications and the possible economic exploitation of the data collected.</p>	<p>Online data management plan (D10.2) completed and available to all partners: http://dmponline.dcc.ac.uk/projects/archaide-horizon-2020-dmp</p> <p>M32 (final release)</p>
6. Evaluating the impact in real conditions	<p>The objective and subjective evaluation of principal system components, performed by means of experiments and user studies developed in real working conditions. The provision of a report outlining an assessment of the effectiveness of each tool component. This report will focus on the evaluation of both the system (functionalities provided) and its practical usability</p>	<p>The first step toward the achievement of this objective has been the realisation of the ArchAIDE system Mock-up, designed for working in real excavation condition. Main steps for evaluating the impact in real working conditions will start on M22 with the beginning of WP8.</p> <p>Month 32(final release)</p>

	evaluated from an end-user point of view.	
7. Establishing the conditions for a successful adoption and exploitation of project outcomes	<p>Since the beginning of the project, a special attention will be addressed to public demonstrations, conferences and awareness rising of key actors in the field. This will include the live demonstration of the system in real excavation settings; the delivery of the tools to sample users even external to the consortium. Setting-up of the dissemination and commercialisation network, including an attractive and regularly updated website and other communication means.</p> <p>Active consideration of standardisation opportunities of the domain.</p>	<p>In the first year of the project, the discussion on the potential exploitation of the outcomes of the project started. In particular, the issue of copyright (re-use of data) and the perspective role and use of the ArchAIDE database were taken into account</p> <p>Month 33 (final release)</p>

1.2 Explanation of the work carried per WP

1.2.1 Work Package 1 – Management

UNIFI, CNR, TAU, UoY, UB, UCO, BARAKA, ELEMENTS, INERA

At present, WP1 is in its 12th month of work, and all the Tasks (all 3 led by UNIFI) are in progress as planned. The main objective of WP1 is to guarantee the correct implementation of the overall management, the monitoring of the project and its activities, and the achievement of the project objectives.

Task 1.1 Technical Management and Quality Control

Leader: UNIFI [other partners: CNR, TAU, UoY, UB, UCO, BARAKA, ELEMENTS, INERA]

Progress planned according to DoA

- Organisation of the Kick-off Meeting.
- Definition of the Management and Technical Board (MTB).
- Organisation of MTB physical Meetings.
- Organisation of MTB video-conference Meetings.
- Creation and maintenance of the Project Risk Register.
- Submission of D1.1 Quality Plan.
- Monitoring of the Project activities.

Actual progress

- Definition of the Management Structure: a special attention was paid to the creation of the Management and Technical Board (MTB) and to the appointment of the members.

- Organisation of the Kick-off Meeting, a Technical Meeting in Barcelona, two Technical Meeting in Pisa and three Skype call MTB Meetings.
 - UNIPI organised the kick off meeting, which was the first official event of the project, and was held in Pisa on 14th-15th June 2016. It has been dedicated to the presentation of the project and to the first MTB meeting. It gave the opportunity to create a good team spirit and to discuss the first steps of the project. Every partner of the consortium attended the meeting.
 - The MTB Skype Call of 7th November 2016 was attended by UNIPI, INERA, CNR, UoY, UCO, BARAKA.
During the Skype Call, it was discussed the development of the website; the gadgets that will be realised during the project; the progress of D10.2; the Internal report procedure as it described in the Quality Plan; the progress of D2.1; the opportunity of engaging professional archaeologists as "external collaborator"; the progress of the amendment procedure; the unpaid leave period of Professor Wolf (TAU).
 - The MTB Skype Call of 11th January 2017 was attended by all the partners.
During the Skype call, it was discussed the progresses of deliverables D2.1 and D1.2, the progress in the achievement of Milestones MS2 and MS3, the changes needed for improving the Internal reporting procedure, the current and upcoming dissemination activities; the progress in the Amendment to the GA procedure.
 - The MTB Technical Meeting of 6th - 7th March 2017 was held in Pisa and attended by all the partners.
The meeting was organised as an open discussion, during which each WP leader made an overview of the progress towards the objectives of the action, illustrating any possible problem and/or risk, with the notification of deviations, their possible impact on the implementation of the action and any measures taken to mitigate the impact of the deviation.
 - The MTB Skype call of 8th May 2016 was attended by all the partners. UoY, UCO and INERA presented any carried out activity and next steps related to WPs 3, 5, 7 and the Reference Database. CNR and TAU explained any carried out activity and next steps related to WPs 4, 6. In particular TAU envisaged their difficulties in achieving D 6.1, due to the lack of enough data for training the neural network classifiers on real pottery data. The risk (R4) was foreseen in the proposal, but the mitigation measures were not suitable. Consequently, TAU shifted its attention to task 6.2 (shape based matching). Finally, MTB decided that once there will be enough data for appearance based similarity training and evaluation, and once there will be a definition, TAU will go back to working on the appearance based matching. UNIPI and UoY explained any carried out activity and next steps related to WPs 9 and 10, with a special focus about the Exploitation Plan (draft version to be discussed during the review meeting in Brussels).
- Submission of D1.1 *Quality Plan*. The *Quality Plan* defines a set of rules for the organization of day-to-day cooperative work within the project, including the procedures to be used, the decision making protocols, the reporting mechanisms, the organization of meetings, the control of information flow, the reliability of the output, and the preparation of documentation for submission to the EC. Furthermore, the document outlines the success criteria for each deliverable, defines the structure of each deliverable, describes the quality review techniques and also defines configuration management procedures and change control. The document serves as a guide for the Project Coordinator (PC), in order to ensure that quality reviews will occur at appropriate points in the project, and will provide the Partners with a concise reference to the project management structure, tasks, responsibilities and reporting requirements.
- Creation and Maintenance of the Project Risk Register.
- Production of three Internal Periodic Reports:
 - M1-6

- M7-9
- M10-12

Task 1.2 Administrative/Financial Management and Reporting

Leader: UNIPI [other partners: CNR, TAU, UoY, UB, UCO, BARAKA, ELEMENTS, INERA]

Progress planned according to DoA

- First transfer of the respective shares of the financial support to the partners.
- Monitoring of the resources employed by the Consortium.

Actual progress

- the Coordinator transferred the first tranche of the respective shares of the financial support received from the EC to all the partners at the beginning of the project.
- Definition of the rules for financial reporting (Section 3.3.2 of D1.1).
- First internal Financial Report at M6.

Task 1.3 Management of knowledge and IPR

Leader: UNIPI [other partners: CNR, TAU, UoY, UB, UCO, BARAKA, ELEMENTS, INERA]

Progress planned according to DoA

- Definition of Knowledge and Intellectual Property Right strategy.

Actual progress

- Discussion at MTB level about the different option for the future exploitation of the project.
- Study of different scenarios.
- Submission of D1.2 IPR and Knowledge Management Plan.

Deliverables Completed:

D1.1 Quality plan M4.

D1.2 IPR and knowledge Management Plan M8.

D1.3 Financial Report 1.

D1.4 Progress Report 1.

1.2.2 Work Package 2 - Methodologies, specification and design

UNIPI, CNR, TAU, UoY, UB, UCO, BARAKA, ELEMENTS, INERA

WP has ended on month 8. The goal of this work package was to define a methodological approach for identifying, gathering and structuring the required data sources in order to design the project database, the data representation schemes, the algorithmic solutions for shape-based recognition/identification and, finally, the overall ICT system specifications.

Task 2.1 Methodologies, scenario definition, user requirement collection and analysis

Leader: UNIPI [other partners: UoY, UB, UCO, BARAKA, ELEMENTS, CNR]

Progress planned according to DoA

- Definition of the typical working scenario in archaeological investigation.
- Specification of the digital supported pipeline for pottery recognition.
- Selection of the main catalogues to be used in the digitisation phase.

Actual progress

- On 15th -16th September 2016, UB team hosted the project second General Meeting in Barcelona followed by a Management and Technical Board Meeting. The meeting was scheduled in order to reach the first Milestone of the project (MS1 Preliminary report about the archaeological and technical specification of the system), especially with regard to:
 - the minimum set of data necessary for making the application works, and
 - the choice of the catalogues needed for developing the application and the test beds.

The partners have agreed to focus on *Terra Sigillata*, Roman amphorae and majolica catalogues. ICT suggestions were revised under the archaeological needs, generating a positive debate where constructive solutions were implemented. A working schedule was also fixed for the next months; a detailed dissemination plan was approved and the editorial board was confirmed. Every partner of the consortium attended the meeting.

- Definition of the working scenario in archaeological investigation and pottery processing with a detailed exposition of the main procedures and objectives in the practice of professional archaeology concerning the study of archaeological ceramics.
- Definition of the Database structure and of the issues related to the Data Preservation.
- Selection of the pottery catalogues. The choice fell on:
 - the amphora types manufactured throughout the Roman world between the late third century BC and the early seventh century AD,
 - the *Terra Sigillata* manufactured throughout the Roman empire,
 - the medieval and post-medieval pottery found in Barcelona and Pisa.

This choice permits to create a first consistent and helpful data pool for archaeologists and to TAU to develop both shape-based and appearance-based similarity algorithms.

Task 2.2 Monitoring the State of the art

Leader: CNR [other partners: INERA, TAU, UNIPI, UB, UCO, UoY]

Progress planned according to DoA

- Monitoring of the state-of-the-art in the use of Computer Graphics in the field of pottery classification.

Actual progress

- Analysis of previous initiatives and papers in order to have an overview of the achieved results, and of the limitation of previous approaches. Since most of the previous works relied on the use of 3D models of the sherds, the outcome of the application of these methods was analysed, showing that an on-site use was impractical, and in any case a strong intervention by user was needed. Nevertheless, the analysis of state-of-the-art provided precious input regarding the main features that could be extracted from the drawings, and used for the description of the typologies and sherds.
- Discussion in regard with the analysis of the state-of-the-art in the archaeological workflow for pottery classification, and on the possible uses of mobile technologies.

Task 2.3 Design of the vectorial and feature-based representation for drawings

Leader: CNR [other partners: TAU, INERA, UoY]

Progress planned according to DoA

- design of the vectorial and feature-based representation of drawings.

Actual progress

- Analysing the state of the art.
- Collecting feedback about the current workflow and needs by the archaeological community.
- Discussing the structure of drawings in the catalogues.
- Based on the outcomes of the above actions, a list of desirable features was discussed, and a shortlist was selected following several criteria (possibility to automatically extract the features, capacity of the feature to describe the typology in a coherent way, etc.).
- Realisation of a prototype of the representation.
- The first results of the automatic extraction of the representation were made available in the project website.

Task 2.4 Detailed functional specification (system architecture, components API) and selection of hardware platforms/devices

Leader: INERA [other partners: CNR, TAU, UNIPI]

Progress planned according to DoA

- Identification of the required functionalities.
- Design of the overall system architecture.
- Analysis of the styles of the user interaction and modalities.

Actual progress

- Realisation and implementation of a Mock-up of the ArchAIDE application
- Discussion about the functional specification of the platform devices. In particular, on everything related to the interaction of user when the feature-based representation of sherd was involved.
- Analysis of the hardware specification about the type of camera (and acquisition environment) needed to provide high quality images for the classification of sherds.

Task 2.5 Selection of the test bed cases

Leader: UNIPI [other partners: UB, UCO, BARAKA, ELEMENTS, CNR]

Progress planned according to DoA

- Selection of relevant case studies to perform a solid assessment of the designed tools according to real archaeological requirements.

Actual progress

- selection of the test bed cases, considering real archaeological investigations (covering both heritage sites and development-led excavations, and surveys), rather than small test cases specifically planned for the purpose of the project.

The selection was based on:

- the evaluation of the site complexity and potential impact with respect to the potentially expected findings (also including the technical constraints related to the digging and documentation);
- how do they fit the archaeological focus and pottery domain selected;
- the availability of documentary sources (iconography, previous archaeological documentation, etc.);
- the tests bed will be carried on the field by BARAKA and by ELEMENTS, whereas UB, UCO and UNIPI will carry test beds on pottery preserved in their potsheds.

Deliverable completed

D2.1 Detailed system specification, including the representation structure for shape data.

Milestone completed:

MS1 Preliminary report about the archaeological and technical specification of the system.

1.2.3 Work Package 3 - Creation of the Application Database

UoY, UNIPI, CNR, UB, UCO, INERA

At present, WP3 is in its 8th month of work. The objective of WP 3 is to carry out the design process of the database, based on the methodologies, specifications and planning carried out within WP2. It is important to ensure the database is both fit for purpose to hold the data upon which the work will be based, and the resulting data created through us of the new technologies. It is equally important that the structure of the database is optimised for archiving, dissemination and re-use of the resulting data, according to open data standards. In order to ensure an interoperable resource, pottery catalogues in a variety of languages will be incorporated using multilingual controlled vocabularies.

Task 3.1: Database design

Leader: UoY [other partners: INERA, CNR, UCO]

Progress planned according to DoA

- Complete database design.
- Coordinate with INERA during implementation phase.
- Respond to any necessary changes in database design during implementation by INERA.

Actual progress

- Taking into account the methodologies, specifications and design planning carried out in WP2, and using the ADS Roman Amphorae database as a point of departure, ADS has completed the database design in partnership with INERA. CNR, UCO and all partners contributed feedback during the kick-off (Pisa) and technical (Barcelona) meetings, along with email correspondence. CNR participated in the initial design of the database. The main contribution was in the definition of the format and typology of data related to the features extracted from the drawings digitised from the catalogues. The features will be used in the context of shape-based classification of sherds.
- A Technical Meeting was organised on 14th December in Pisa. Catalogues digitisation, database structure, design of vectorial and feature-based representation for drawings were discussed.
- Database design completed and changes made as required during implementation.
- Coordinated with INERA as the design moved into the implementation phase.

Task 3.2: Incorporation of multilingual vocabularies

Leader: UoY [other partners: UB, UNIPI, UCO]

Progress planned according to DoA

- Scope potential for using the Getty Art and Architecture thesaurus (AAT) as a neutral spine for creating multi-lingual mappings pottery characteristics, using tools developed by the ARIADNE project.
- Begin creating multilingual vocabularies of pottery characteristics by UoY, UB, UNIPI and UCO.
- If found suitable after scoping exercise, implement use of Getty Art and Architecture thesaurus as a neutral spine for creating multi-lingual mappings pottery characteristics, using tools developed by the ARIADNE project.

- Coordinate creation of multilingual vocabulary mappings of pottery characteristics by UoY, UB, UNIPI and UCO.

Actual progress

- Upon completion of the database design, work began on incorporating multilingual vocabularies.
- Scoping of potential for using the Getty Art and Architecture thesaurus as a neutral spine for creating multi-lingual mappings pottery characteristics was completed, with positive results.
- Mapping of multilingual vocabularies of pottery characteristics to the Getty Art and Architecture thesaurus completed by UoY, UB, UNIPI and UCO.
- Mappings transformed to XSLT by UoY and delivered to INERA in JSON format for implementation.

Task 3.3: Database implementation

Leader: INERA [other partners: UoY]

Progress planned according to DoA

- INERA to begin implementation of the comparative database, based on the design completed by ADS.
- INERA to begin implementation of the multi-lingual vocabularies created in Task 3.2.
- UoY to work closely together INERA during implementation of the comparative database, and during the initial population of the database.

Actual progress

- INERA has completed implementation of the comparative database, based on the design completed by ADS, and partners are carrying out testing.
- Population of the database with comparative data is underway.
- The implementation of the multi-lingual vocabulary functionality has been implemented.

Milestone completed:

MS2 First Version of the database ready for being populated.

1.2.4 Work Package 4 - Technologies for the digitization of catalogues

CNR, UNIPI, TAU, UCO, INERA

At present, WP4 has reached its 8th month of work. The goal of this Work Package is to study and test technologies to transform the paper catalogues describing the pottery classes in a digitized version. The items to be taken into account in this digitization effort are the ones defined in WP2. The goal here is to design a pipeline and the relative technologies to transform mostly printed content into a structured digital representation.

Task 4.1: Technologies for the digitization of catalogues

Leader: CNR [other partners: INERA, UCO, UNIPI]

Progress planned according to DoA

- Analysis of the selected paper catalogues.
- Testing of available OCR tools.
- Implementation of an ad-hoc solution.

Actual progress

- Different solutions for OCR have been analysed, in order to analyse flexibility, amount of languages supported, extendibility. The analysis was brought on by ISTI-CNR, UNIPI and INERA.

- *Tesseract* set of tools was selected for the implementation in the context of the interface for database population. It was selected since its JavaScript-based implementation made it easy to integrate in a web-interface, and due to the fact that several languages, including Italian, English, Spanish, and French, are supported.
- Two different implementation (created by ISTI-CNR, with the support of INERA) of the OCR-based tool have been proposed. The first implementation is a basic OCR, where the selected text is automatically read and made available.
- The second OCR implementation (created by ISTI-CNR, with the support of INERA) aims at automatically filling the database structure describing by analysing the structure of the description. The “*Conspectus*” catalogue was selected to implement and test this solution.
- The initial implementations of the OCR systems have been presented and discussed during the Pisa meeting in March 2017. Several issues related to the interface and data exchange have been discussed. All the partners took part in the discussion.
- The basic OCR tool, which aims at extracting the text from generic catalogues, has been improved, and its integration in the context of the Database interface proposed by INERA has started. The mechanism to select and assign text to the different fields describing the elements of the database was defined. The work was brought on by CNR, INERA and UNIPI.
- The advanced OCR implementation (that aims at automatically filling the database structure describing by analysing the structure of the description) is currently under finalization for the part related to the automatic production of a JSON describing every single type. The “*Conspectus*” catalogue was selected to implement and test this solution. The work was brought on by CNR and INERA.
- A further extension of the tools to (semi-)automatically extract the depictions from the images is currently in its design phase.

Task 4.2: Detection extraction and vectorialization of drawings from digitized pages

Leader: CNR [other partners: TAU, INERA]

Progress planned according to DoA

- Produce a vectorial representation of drawings.
- Build a tool to automatically extract the representation.

Actual progress

- The structure of the drawings in the context of several catalogues has been analysed, in order to select the geometrical features that can be reasonably extracted in all cases. This led to the definition of a set of features that include: internal and external profile of the body, internal and external profile of the handle (if present), rim and base points, scale factor, mouth radius. The analysis and selection was made by all the partners involved in the WP.
- An automatic application that is able to analyse a scanned version of the drawing has been implemented. The application (implemented by ISTI-CNR) is able to automatically extract the above mentioned features, and save the information in the context of an .svg file.
- In addition to the creation of the .svg structure, a 3D representation of the drawing is created automatically by taking into account also the axis of revolution (indicated in the drawings as well). The application (implemented by ISTI-CNR) is currently able to extract both the body and the handles. The creation of a unique model is currently under implementation.
- The development of the drawing extraction tool is currently integrated with the work of WP6.2 (in collaboration with TAU), that aims at creating an automatic classification system based on shapes. Both the .svg files and the 3D models will be used in the creation and training of the classification system.

- The semi-automatic tool for the extraction of geometric features from drawings has been presented and validated during the Pisa meeting in March 2017. All the partners took part in the discussion.
- A further automatic step of the digitization, aimed at producing a single 3D model from the base and handles model, has been added.
- The development of the drawing extraction tool is currently integrated with the work of WP6.2 (in collaboration with TAU), that aims at creating an automatic classification system based on shapes. An initial testing set of .svg files and the 3D models has been created and shared with TAU for the initial development and testing of the shape-based classification system.
- A paper regarding the full pipeline for 3D reconstruction and virtual fragmentation of 3D models has been submitted to the ICDAR 2017 conference. The submission is currently under review.

Task 4.3: Preliminary validation of digitization technologies on sample datasets

Leader: CNR [other partners: UCO, UNIPI]

Progress planned according to DoA

- Test and validate the OCR tools.
- Test and validate the drawing digitization tool.

Actual progress

- Initial tests have been made on both the OCR (mainly on the “Conspectus” catalogue) and on the automatic digitization of drawings (mainly drawings coming from the ADS Amphorae database). This work has been supported by UNIPI and UCO.
- The implementation of the automatic OCR (on the “Conspectus” catalogue) is under testing: a set of initial data has been produced.
- The 3D generation tools are currently under use for the initial population of the ArchAIDE database.

1.2.5 Work Package 5 – Population of the Database

UCO, UNIPI, CNR, UoY, UB, BARAKA, ELEMENTS

At present, WP5 is in its 4th month of work. The main goal of this work package is to populate the newly designed database with data using the technologies developed in WP4.

Task 5.1: Search and selection of paper catalogues

Leader: UCO [other partners: UB, CNR, BARAKA, ELEMENTS, UNIPI]

Progress planned according to DoA

- Besides data already accessible in external databases, the ArchAIDE database will also be filled with new data deriving from paper catalogues.
- The partners will finalize the search and selection of paper catalogues according to the chosen pottery classes and the requirements of the project as well as the preliminary information produced by WP4.

Actual progress

- Taking into account the specifications identified in WP2, agreement on pottery classes Amphorae and *Terra Sigillata* and medieval e postmedieval majolica.
- Selection of for classes important catalogues.
- UNIPI with UB and UCO organised the lists of Roman amphorae and *Terra Sigillata Italica* form types harmonizing names and different typologies coming from several different studies.
- Making catalogues available for CNR to text OCR and 3D tool.

Task 5.2: Automatic digitization and OCR

Leader: UCO [other partners: UB, CNR, UNIPI]

Progress planned according to DoA

- Digitization of the chosen paper catalogues.
- Information extraction approach to reduce long textual fields to key value pairs.

Actual progress

- Digitization of catalogues made by UCO, UB and UNIPI:
 - AAVV, *Conspectus Formarum Terrae Sigillatae italico modo confectae*;
 - BERTI F., *Storia delle Ceramiche di Montelupo* (4 volumes);
 - DRAGENDORFF H., *Terra Sigillata*;
 - FORNACIARI A., *La sostanza delle Forme: Morfologia e cronotipologia della Maiolica di Montelupo Fiorentino*;
 - GEMPELER R.D., *Elephantine X, Die Keramik römischer bis früh-arabischer Zeit*;
 - HAYES J. W., *Late Roman Pottery*;
 - MEDRI M., *Terra sigillata tardo italica decorate*;
 - MAYET F., *Les Céramiques Sigillées Hispaniques*
 - OXÉ A., COMFORTH., KENRICK PH., *Corpus Vasorum Arretinorum. A Catalogue of the Signatures, Shapes and Chronology of Italian Sigillata*.
- First tests of the OCR-tool created by CNR (*Conspectus*).
- Selection of next catalogue for CNR to train the tool .
- Research in field of text mining and information extraction.

Task 5.3: Database population

Leader: UCO [other partners: UB, UoY, UNIPI]

Progress planned according to DoA

- The partners will insert the data produced in the project database
- A workflow from digitization to population has to be developed so partners can use it in early stages of the project

Actual progress

- First manual population of database to see functionality and bugs
- Defining a workflow for populating the database
- Taking picture of potsherds to feed the Database:
 - BARAKA has been working in order to obtain the cooperation of several Spanish research teams studying *Terra Sigillata Hispanica*. In this sense, BARAKA has achieved the commitment of the following Research Teams:
 - Research Team of Consorcio de Mérida (Mérida, Badajoz, Spain): Ancient Roman City of Emerita Augusta, in the southwest Spain, whose archaeological contexts are specially indicated to do the testbeds of Terra Sigillata Hispanica due to their outstanding formal variety;
 - Research Team of Medellín archaeological site (Medellín, Badajoz, Spain): Ancient Roman Site of Colonia Metellinum, in the southwest Spain;
 - Research Team of Cáceres El Viejo archaeological site (Cáceres, Spain): Ancient Roman Camp of Castra Caecilia, in the southwest Spain;
 - Research Team of Regina archaeological site (Casas de Reina, Badajoz, Spain): Ancient Roman City of Regina, in the southwest Spain;

- Research Team of Los Villares archaeological site (Andújar, Jaén, Spain): Ancient TSH production centre in the South of Spain;
- ELEMENTS is working on recording different sets of archaeological sherds. As a first step, the archaeological sites from Mallorca in which Roman amphorae and *Terra Sigillata* types have been recovered are identified. Afterwards, the archaeological materials that are stored in the premises of the Museum of Mallorca have been specifically located. In this moment, Elements is recording by photography the set of materials from the following sites:
 - Cabrera V (Palma de Mallorca, Spain);
 - Colònia de Sant Jordi (Palma de Mallorca, Spain);
 - Porto Pi (Palma de Mallorca, Spain);
 - Roman city of Pollentia (Alcúdia, Mallorca, Spain);
 - Urban excavation from c/ Colon, 5 (Palma de Mallorca, Spain);
 - Excavation from Andratx (Mallorca, Spain).

Furthermore, nowadays other archaeological sites where Medieval pottery has been recovered are also being identified and the materials specifically located in the premises of the Museum of Mallorca in order to be recorded by photography as well. Simultaneously, the typology of these materials that are already published are being digitized.

Additionally, a partnership agreement has been achieved with the Cultural Heritage Department of the city hall of Andratx (Mallorca) in order to record by photography the archaeological materials recovered around this municipality.

- UB is working on Roman amphorae and *Terra Sigillata*, and Medieval and Postmedieval pottery present in municipal archaeological warehouses. The Roman amphorae that have been already documented (52 different types), came from the following survey excavations from the city of Barcelona:
 - Carrer Sostinent Navarro;
 - Mercat de Santa Caterina;
 - Plaça del Rei;
 - Zona Franca;
 - Arxiu Administratiu;
 - Carrer Sant Honorat;
 - Carrer Lledó.

In the case of *Terra Sigillata*, UB will have the opportunity to take pictures of different types from the following archaeological sites:

- Clunia;
- Baetulo;
- Tarraco;
- Emporion;
- El Vilarenc;
- Barcino;
- Andújar.

For Medieval and Postmedieval pottery different assemblages from the museum support center of the Museu d'Història de Barcelona will be available for the population of the database.

- UNIPI is working on Roman amphorae, *Terra Sigillata Italica* (TSI) and Medieval and Postmedieval Maiolica of Montelupo Fiorentino, from the following excavations:
 - Piazza Duomo (Pisa, Italy);
 - Via Galluppi (Pisa, Italy);
 - Piazza Aranci (Massa, Italy);
 - Area archeologica Massaciucoli romana (Massarosa, Lucca, Italy);
 - Vada Volterrana (Rosignano Marittimo, Livorno, Italy).

UNIPI has achieved the cooperation of other archaeological research group, in order to take picture from the following archaeological excavations:

- Colosseo (Rome, Italy);
- Fori Imperiali (Rome, Italy);
- Development-led excavation (Rome, Italy);
- Spoletino (Viterbo, Italy);
- Scoppieto (Terni, Italy.)

UNIPI has achieved the opportunity of taking pictures of *Terra Sigillata Italica* form types conserved at:

- Museo archeologico Gaio Cilnio Mecenate, Arezzo
- UCO is in dialogue with the Römisch-Germanisches Museum Cologne about possible photo campaigns. Due to the restauration and closing of the museum there is no further development made.

Furthermore, colleagues from the CERAMALEX project try to get photos of amphora types from colleagues in Egypt, coming from following archaeological sites:

- Alexandria (Egypt);
- Schedia (Egypt)

Task 5.4: Normalisation of the CERAMALEX data on Hellenistic and Roman Pottery in Alexandria

Leader: UCO

Progress planned according to DoA

- Normalization of the CERAMALEX data is crucial for import in the project database.
- A mapping of the different database fields will be undertaken to create a mapping tool that can be used for other highly structured data as well.

Actual progress

- Export of CERAMALEX-ERD and comparison to ArchAIDE-data model.
- Mapping of CERAMALEX fields with the ArchAIDE fields.
- Mapping of controlled vocabulary - multilingual vocabulary is created by UNIPI, UYO, UB and UCO.

1.2.6 Work Package 6 - Shape and image-based similarity search and retrieval

TAU, UNIPI, CNR, INERA

At present, WP6 is in its 8th month of work. The task specific work for WP6 has focused so far on building synthetic training data of 3D vessels as well as simulating breaking them and obtaining matching sherds. This data will be used in order to develop the algorithms necessary for matching sherds with catalogue-based drawings of vessels.

Task 6.1: Appearance-based search and retrieval over a database of shapes

Leader: TAU [other partners: CNR]

Progress planned according to DoA

- Algorithms for pairwise similarity between ceramics based on appearance (software).

Actual progress

- While the database of ceramics was populated (WP5), an initial overview of the data was done by TAU to identify the sort of similarity to be detected as part of this task. When doing so, it was discovered that there are some problems with this data. Specifically:

- **the data consists of scans of entire catalogue pages** – not dedicated images of the pottery samples cropped from the rest of the page;
- **The data is unlabelled** – a critical component for training a machine learning algorithm is to actually have the input data classified and labelled so that we can train the algorithm to recognize similar images as belonging to this class. Unfortunately, the current data is not labelled in any automatically consumable way; annotations are sometimes included in big paragraphs of text, but not in a machine consumable manner;
- **there is no unified archaeological definition for “appearance-based similarity”** – In order to determine which pieces of ceramics could be considered as belonging to the same class, there must be some definition on when is the appearance difference is actually acceptable (and is considered as the same class) and when it’s too big (and therefore considered of different classes). However, in recent discussions it was raised that there’s no archaeological consensus on how to measure similarity as:
 - in some cases, it’s defined mainly by colour (and patterns may change);
 - in others it’s defined by the pattern;
 - and more.

For this reason, it’s currently not possible to develop a similarity algorithm – since there’s no definition of what is considered similar;
- **Lacking data within classes** –as described in the work package “the similarity would be learned from examples that are tagged as similar or not-similar”. However, this requires **many samples per class**, and currently some catalogues only present one canonical image per class, thus preventing learning the similarity.

These problems were identified and surfaced to the partners. Some of the solutions suggested were to shift the focus toward engraved signatures and other appearance based clues. The TAU team can accommodate such new data very efficiently.

Meanwhile, generic appearance based networks were trained by a TAU student, which can be adapted to the new appearance data given training samples.

Task 6.2: Shape-based search and retrieval over a database of shapes

Leader: TAU [other partners: CNR]

Progress planned according to DoA

- Algorithms for pairwise similarity between ceramics based on shape and appearance should only be ready by M16, listed here since we had progress in this task.

Actual progress

- Task 6.2 poses similar challenges to task 6.1, but the reliance on shape only and not on texture and appearance, allows much more freedom in gaining the data. Therefore, most of the work was indeed acquiring data, by synthetically generating data. The idea is to:
 - import geometric features (inner and outer base profiles) as imported from pottery drawings by using the software developed in task 4.2;
 - generating full 3D models for each class of pottery, by rotating the profiles from the previous step;
 - using 3D algorithms to virtually break each model to virtual sherds, and repeating this to obtain many different sherds for each class;
 - aligning the sherds in front a virtual camera, like an archaeologist would do, to capture “pictures” of sherds;
 - extracting the shape outline from each such image, to obtain the final shape data.

During our work period, we completed most of these efforts. There’s still some work on completing the implementation of the alignment algorithm, and after that we expect starting to generate the synthetic sherds to have data for training the machine learning algorithms.

- Also, in a joint work with CNR, we submitted a paper for ICDAR 2017 regarding the 3D generation pipeline. The paper is currently under review.

1.2.7 Work Package 7 - The mobile tool and Front-end Desktop Application

INERA, UNIPI, CNR, UoY, UB, BARAKA, ELEMENTS

At present, WP6 is in its 4th month of work. Currently, only Tasks 7.1 and 7.3 are active. The goal of this WP is provide a usable tool following the requirements coming from WP2, integrating the search and retrieval mechanism defined and developed in WP6.

Task 7.1: Detailed design of the tool

Leader: INERA [other partners: CNR, UNIPI, BARAK, ELEMENTS; UB]

Progress planned according to DoA

- Animate and coordinate the discussion about the tool functionalities and the scenarios where archaeologists would use it.
- Define the hardware and software framework to be used in the implementation of the tool.
- Define detailed functional specification of the tool.

Actual progress

- Completed the analysis of the potential scenarios where the mobile tool would be helpful. The activity is carried on using the User Centred Design methodology and aims at producing a detailed definition of the user experiences and the requirements for the mobile and desktop tool.
- Defined the technical architecture for the mobile tool and desktop tool.
- Almost completed the definition of a detailed mock-up of the mobile tool.
- First draft of the detailed functional specification of the tool.

Task 7.3: Implementation of the mobile application

Leader: INERA [other partners: CNR]

Progress planned according to DoA

- Definition the communication protocol between mobile tool, repository database and search tools.
- Implement the Result database (i.e. server repository) and communication API.

Actual progress

- Started the implementation of the core components of the mobile tool (communication, user authentication, 2D image acquisition, etc.).
- Implemented the Reference Database API for the retrieval of the ceramic types from the mobile tool through the search engine produced in WP6.
- Started the implementation of the “Result database” (server database supporting the mobile tool).
- Started the implementation API interface enabling the communication between mobile and desktop.

1.2.8 Work Package 8 - Test and assessment of the overall system on application scenarios

WP 8 will begin at month 22.

1.2.9 Work Package 9 - Communication/Public engagement/Innovation

UNIPI, CNR, TAU, UoY, UB, UCO, BARAKA, ELEMENTS, INERA

At present, WP9 is in its 12th month of work, and all the Tasks are in progress as planned. The main objective of this WP is communicate the implementation of the project, in order to maximise its impact beyond the project community.

Task 9.1: Communication Management

Leader: UNIPI [other partners: all]

Progress planned according to DoA

- Definition of Communication Plan.
- Activities of communication management.
- Actions about media presence and relations.

Actual progress

- The Kick-off meeting has been organised on 14th -15th June 2016 in Pisa (Gipsoteca di Arte Antica, Piazza San Paolo all'Orto 20), by UNIPI and has been articulated in a public morning section, and a section reserved to the consortium partners (WP1).
- The Communication Plan was defined during M2 and M3 (D9.1 – 31.08.2016). It has involved all the partners of the consortium and described aims and strategies about the communication activities that will be carried out during the whole project. The Plan has been written simultaneously with the Dissemination Plan (D10.1) and it is closely related to it. In the Communication Plan has been underlined the importance of the role of the communication during all the phases of the project, the importance of a transparent way for describing each step of the work, and the choice of using a "visual storytelling" strategy to engage an as large as possible audience. The final editing of the Communication Plan has been made by UNIPI, INERA and UoY.

The main goals are:

- making know the existence and the aims of the project;
- producing an expectation about the final products during the time of the project;
- reaching a wider audience, beyond the scientific and archaeological community;
- involving the general public through a smart and friendly use of language;
- explaining to the archaeological community the benefits connected with the introduction of ArchAIDE tool and involving them in testing it;
- building a strong community for carrying on the communication activities during the exploitation phase and after the end of the project.
- Creation of the Editorial Board. During the General Meeting in Barcelona (15th -16th September 2016) the creation of the Editorial Board was approved. The board is composed by one representative from each partner and two technical members (a graphic designer and a filmmaker), and led by UNIPI. Members and activities of the Editorial Board are described in D 9.2.
- Media presence:
 - a first press releases have been written after the signature of the Grant Agreement by UNIPI (January 2016);
 - a second press releases has been realised for the Kick-off meeting held in Pisa on 14th -15th June 2016;
 - after the kick-off meeting, 9 articles item were published by Italian newspaper and websites;
 - a press conference for the public presentation of the project and the website has been made in Rome on December 2, 2016.
 - in occasion of the press conference, a press releases have been written in 5 languages (Italian, German, English, Spanish and Catalan) by UNIPI, UCO, UoY, BARAKA and UB.
 - a downloadable Press Kit is available on the "Press" page of ArchAIDE website (<http://www.archaide.eu/press>);
 - after the press conference, 32 articles item were published on newspapers and web media.

Table of published news about ArchAIDE on newspapers and web media:

Date	Media	Country	Readers for month
------	-------	---------	-------------------

17/04/2017	Ultima Hora	Spain	232.000
11/12/2016	Cittadini di Twitter	Italy	920
09/12/2016	Cinque Colonne Magazine	Italy	1.700
06/12/2016	La Discussione	Italy	1.600
05/12/2016	Koeln Nachrichten	Germany	20.000
05/12/2016	LPL News 24	Italy	1.200
05/12/2016	Universitat zu Koeln	Germany	50.000
05/12/2016	IDW	Germany	35.000
05/12/2016	InToscana.it	Italy	41.900
04/12/2016	Nova 24	Italy	193960
03/12/2016	El Diario	Spain	164.833
03/12/2016	Arquelogia y patrimonio virtual	Spain	17.418
02/12/2016	Digitalmeetsculture.net	International website	10.800
02/12/2016	Qui news pisa	Italy	6.500
02/12/2016	Arte Magazine	Italy	7.100
02/12/2016	Gonews.it	Italy	50.400
02/12/2016	Toscana24– Il Sole 24 ore	Italy	1.400.000
02/12/2016	unipi.it	Italy	121.942
02/12/2016	askanews.it	Italy	71.700
02/12/2016	meteoweb.it	Italy	199.300
01/12/2016	EfeFuturo	Spain	5.045
01/12/2016	La Bitacora de Jenri	Spain	4.500
01/12/2016	Tualbacete	Spain	62.579
01/12/2016	Diario Lanza	Spain	156.962
01/12/2016	La Tribuna de Ciudad Real	Spain	175.968
30/11/2016	Pileta de Prehistoria	Spain	4.500
30/11/2016	Objetivo Castilla la Mancha	Spain	18.382
30/11/2016	Archeomatica	Italy	3.300
30/11/2016	Taracon digital.es	Spain	3.300
23/11/2016	Science 4 Data	International website	No data found
30/11/2016	cml24.es	Spain	618.284
30/11/2016	Mibact.it	Italy	286.000
07/08/2016	ResearchItaly	Italy	4.700
18/05/2016	Miningfulstudio blog	Italy	No data found
22/01/2016	Il Giornale d'Italia	Italy	14.500
21/01/2016	Meteoweb.eu	Italy	145.000
21/01/2016	Pisa24	Italy	No data found
21/01/2016	PisaInforma.it	Italy	No data found
21/01/2016	Libero 24X7	Italy	No data found
21/01/2016	La Nazione	Italy	76.621
21/01/2016	QuiNewsPisa.it	Italy	1.000

Task 9.2: Promotional kit

Leader: UNIFI

Progress planned according to DoA

- Design of the logo and visual identity.
- Trademarking of the logo at the EU Office for Harmonization in the Internal Market (OHIM).
- Definition of the promotional kit.

Actual progress

- Design of visual identity (completed by UNIFI):
 - The logo of the project has been designed (M1-3);
 - The visual identity of the project around the logo has been created.
 Logo and visual identity are described in D9.1 Annex 1.
- Promotional kit (completed by UNIFI): a line of products has been designed to promote the project and its aims to a wider public:
 - an A3 leaflet containing general information and available in the five languages has been produced in 5.000 copies (description in D9.2 §5.1). The leaflet is also available in .PDF format for downloading from the “Press” page of the project website. Contents have been realized by UNIFI, UoY, UCO, UB, TAU and BARAKA;
 - customized folders have been produced;
 - customized USB drive, representing an amphora with ArchAIDE logo, has been produced (description in D9.2 §5.2);
 - customized pens, pins and magnets have been designed but not produced yet.

Task 9.3: Web communication

Leader: UNIFI [other partners: all]

Progress planned according to DoA

- Creation and testing of the internal communication network (intranet website).
- Design and creation of the public website.
- Population of the website with contents.
- Creation of infographics and short interviews.
- Creation of a newsletter.

Actual progress

- An internal website has been defined. The intranet is a private website accessible only by team members as registered users. The first structure of the intranet started at the beginning of M3 and has been tested by all partner in the consecutive two months. The site is used as internal network to share documents, drafts, comments, etc. It is also a place of discussion about issues of each WP or concerning the project as a whole. The structure is described in D9.1 and D9.2. The construction of the intranet has been carried out by INERA in cooperation with UNIFI.
- A public website (www.archaide.eu) has been realised:
 - since M4 the conceptual and technical structure of the public website was built (D9.1 §4.1). The sections and the structure of the home page were defined as a work in progress, in order to convey the concepts described in D9.1 with a focus on the visual communication (by UNIFI and INERA);
 - UNIFI carried out the general sections and all partner contributed to the section “partner”, “team”, “blog&news”, “events”, “media” and “press”;
 - the website is on-line since November 28, 2016 (<http://www.archaide.eu>);
 - a calendar of publication on the website blog has been scheduled since 30.11.2016 until 21.12.2017; all partner are involved to create contents three times a year.

Date	Title	Author	Team
16.11.2016	Is automatic pottery recognition possible?	Gabriele Gattiglia	UNIFI

23.11.2016	Deep impact (through the barricades)	Gabriele Gattiglia	UNIFI
15.12.2016	Thinking spatially	Tim Evans	ADS – UoY
03.03.2017	A very useful tool for the archaeologist	Luis Alejandro García García	BARAKA
07.03.2017	From drawing to the 3D representation	Matteo Dellepiane	CNR
20.03.2017	How to train ArchAIDE	Eva Mighuel Gascon	UB
20.04.2017	Virtual sherd for better learning	Barak Itkin	TAU
05.05.2017	Collecting data and improving resolution	Llorenç Villa	ELEMENTS
18.05.2017	ArchAIDE@CAA	Michael Remmy	UCO

- after March Technical Meeting, the website was updated with a new structure of the home page with more emphasis on the video section (completed by UNIFI and INERA). Since November 28, 2016 to May 31, 2017, the website involved 1.029 users with 5.992 visualisations, mainly from Italy, Spain, United Kingdom, Germany and United States;
- in M12 a new section dedicated to “Associates” has been added to involve the archaeological community (<http://www.archaide.eu/associates>). The page gives the opportunity to collaborate with ArchAIDE project as associates, answering a questionnaire and submitting an application. The system allows creating a personal page for each associate and a dedicated space into the intranet. (completed by UoY, UNIFI, INERA).
- Social media communication:
 - opening Social networks profiles: Facebook and Twitter (M2-M3), Instagram, Slideshare and Youtube (M4-M5). Details about social media presence are described in D9.1, §5.5 and D9.2 §4. (completed by UNIFI);
 - a calendar of publication on the social media was planned since 1.10.2016 to 30.06.2017; all partners are involved in creating two contents for month (completed by ALL partners;)
 - periodical monitoring has been made through the standard social media metrics (insights, followers, “likes”, etc.).

Social network	Account	N. post	M6 (28.11.2016)	M12 (31.05.2017)	Difference
FACEBOOK	Archaide project (@archaideproject)	151	Likes on the page: 575	Likes on the page: 1158	+101,4%
TWITTER	@ArchAIDEproject	153	Followers: 139	Followers: 258	+85,6%
INSTAGRAM	@archaide_project	67	Followers: 90	Followers: 296	+228%
YOUTUBE	ArchAIDE project	23	Visualisations (1/07/16 – 28/11/16): 229	Visualisations (29.11.16– 31.05.17): 1609	+602,62%

- Video communication: the creation of video contents has started at the beginning of the project. The activities have been carried out by UNIPI. All the video contents have been published on the web site (<http://www.archaide.eu>), the ArchAIDE YouTube channel (<https://www.youtube.com/channel/UCbQxLFXGaGRtXIBuCqjkNNG>), and the social networks:
 - 10 clips (30-60 sec) with short interviews to each partner have been produced during the kick-off meeting;
 - the short interviews are also part of a longer video (240 sec.) showing the kick-off meeting (https://www.youtube.com/watch?v=-tbG0wuK_xU);
 - a time-lapse video (30 sec) of the kick-off meeting has also been produced (<https://www.youtube.com/watch?v=Vj5NwyZiTxE>);
 - a first animated infographic based on the logo of the project has been created (<https://www.youtube.com/watch?v=yDbO3Eww8wU>);
 - 2 clips (30-60 sec) with interviews to each partner have been produced during the Technical Meeting on 14th December in Pisa (<https://www.youtube.com/watch?v=JXRj6WDAxWk>), (<https://www.youtube.com/watch?v=yI-r3M8rMsE>);
 - a second animated infographic with the main steps of ArchAIDE tool has been created (<https://www.youtube.com/watch?v=xn8P7pTwElw>);
 - 9 clips (30-60 sec) with interviews to each partner have been produced during the Technical Meeting on 7th-8th March in Pisa; 4 of them are completed (<https://youtu.be/1MxqXf2DeLY>; <https://youtu.be/R5WJOGJDI6Y>; <https://youtu.be/jwZDi8IINXM>; <https://youtu.be/obNK1D-AScY>), 5 are in post-production phase.

Task 9.4 Multiplier events

Leader: UNIPI [other partners: all]

Progress planned according to DoA

- Individuation of time and location for the first multiplier event.

Actual progress

- The first multiplier event has been scheduled for 7 December 2017 in York (completed by UoY and UNIPI).

Task 9.5: Production of documentary video

Leader: UNIPI [other partners: all]

Progress planned according to DoA

- Creation first video contents.

Actual progress

- All video contents have been produced to be used also in the documentary video.
- A plan for self-made clips filmed by partners has been scheduled during the last Technical Meeting on March 2017 in Pisa. (completed by UNIPI with collaboration of ALL partner).
- In May 2017, it has been realized the first self-made short video at CNR-ISTI in Pisa to show the current state of the art in the population of the reference database. About 20 clips have been filmed with interviews, details and close-ups, easy actions and establishing shots. The filmed material has been edited and post-produced by the communication team in order to provide a short self-made

video to be uploaded on the project website (<https://youtu.be/lvT1SViTIL4>) (completed by UNIPI and CNR).

- It has been elaborated a Guideline to allow a harmonized self-made production of short video clips showing the achievement of each partner specific tasks and their developments within the project. A dedicated section in intranet website has been created for these materials (completed by UNIPI).

Deliverables Completed:

D9.1 Communication Plan (M3)

D9.2 Website and promotional kit Plan (M6)

1.2.10 Work Package 10 – Archiving and Dissemination

UoY, UNIPI, CNR, TAU, UB, UCO, BARAKA, ELEMENTS, INERA

At present, WP9 is in its 12th month of work. The objective of this work package is to archive, disseminate and promote the results of the project, to follow best practice with regard to open data use and re-use, and to make the project collaborative and transparent to researchers in the archaeological domain and beyond.

Task 10.1: Data Management

Leader: UoY [other partners: UNIPI]

Progress planned according to DoA

- Provide guidance to partners in various areas of data management now that the project is underway, including consultation of D10.2 *Data management plan* which is available at <http://dmponline.dcc.ac.uk/projects/archaide-horizon-2020-dmp>.

Actual progress

- Advice on data management has been given by ADS during the course of the Kick-off Meeting (Pisa); the first technical meeting (Barcelona), the second technical meeting (Pisa) and support is been ongoing via the ArchAIDE shared workspace. D10.2 *Data management plan* was completed, in the form of an online document written using the templates within the 'DMPonline' tool: part of the the Open Research Data Pilot (ORD) funded under Horizon 2020. The ArchAIDE DMP is live at: <https://dmponline.dcc.ac.uk/projects/archaide-horizon-2020-dmp>.

Task 10.2: Data archiving

Leader: UoY [other partners: UNIPI, UB, INERA]

Progress planned according to DoA

- Planning for archiving is ongoing and will be carried out during the final year of the project.

Actual progress

- Planning for archiving is ongoing and will be carried out during the final year of the project.

Task 10.3: Demonstration to the research community

Leader: UNIPI [other partners: CNR, TAU, UNIPI, UB, UCO, INERA, BARAKA, ELEMENTS]

Progress planned according to DoA

- Training Open Days will be planned further into the project when it is possible to demonstrate results of the project.

Actual progress

- Training Open Days will be planned further into the project when it is possible to demonstrate results of the project.

Task 10.4: Dissemination and promotion

Leader: UoY [other partners: CNR, TAU, UNIFI, UB, UCO, INERA, BARAKA, ELEMENTS]

Progress planned according to DoA

- UoY, UNIFI and UCO submitted a poster to introduce the project at the CAA conference in Atlanta, USA, March 14-16, 2017 which was accepted. The poster is titled *Archaeological Automatic Interpretation and Documentation of cEramics (ArchAIDE)*. Holly Wright from UoY and Michael Remmy from UCO will present the poster.
- UoY and UNIFI session proposal for the EAA conference in Maastricht, 30 August to 3 September, 2017 has been accepted. The session is titled *Automation in artefact recognition. Perspectives and challenges in archaeological practice*. The proposal will be promoted as part of the conference call for papers.
- UoY will work with UNIFI and INERA to create the online infrastructure within the ArchAIDE website and workspace to build a community of ArchAIDE Associates comprised of the target dissemination audience set out in D10.1.
- UoY will consult with partners about other dissemination venues, and build a list of completed, pending and potential activities in order to connect with potential ArchAIDE Associates.

Actual progress

- Holly Wright from UoY and Michael Remmy from UCO presented the poster to introduce the project at the CAA conference in Atlanta, USA, March 14-16, 2017. Sixteen people signed up to potentially participate in the project and become ArchAIDE Associates.
- UNIFI presented a paper to AIUCD in Rome, 26-28 January, 2017.
- UoY and UNIFI session proposal for the EAA conference in Maastricht, 30 August to 3 September, 2017 has been accepted. The proposal was promoted and four papers were accepted. The session will be promoted using ArchAIDE communication channels prior to the conference.
- The abstract submitted by BU to EMAC in Bordeaux, 6-9 September, 2017 was accepted as a poster.
- An abstract was submitted by CNR and TAL to ICDAR in Kyoto 9-15 September, 2017.
- UoY worked with UNIFI and INERA to create the online infrastructure within the ArchAIDE website and workspace, to build a community of ArchAIDE Associates, and names captured at CAA and other venues were contacted and invited to join the community.
- UoY consulted with partners about other dissemination venues, and built a list of completed, pending and potential activities in order to connect with potential ArchAIDE Associates.

Task 10.5 Title: Demonstration activity oriented to the market

Leader: INERA [other partners: UNIFI, CNR, BARAKA, ELEMENTS]

Progress planned according to DoA

- Demonstration activities will be carried out in the final year of the project, but planning discussions have begun.

Actual progress

- Demonstration activities will be carried out in the final year of the project, but planning discussions have begun.

Task 10.6: Exploitation

Leader: CNR [other partners: INERA, UNIPI, BARRAKA, ELEMENTS]

Progress planned according to DoA

- Exploitation activities will be carried out in the final year of the project, but planning discussions have begun.

Actual progress

- CNR has created a draft exploitation plan. (preliminary version of D10.5, to be delivered at M33). The current content of this doc is briefly presented here, see Subsection 2.

Deliverables Completed:

D10.1 Dissemination Plan M3

D10.2 Data management Plan M6

1.3 Impact

We consider the information on how ArchAIDE project will contribute to the expected impacts still relevant.

Expected Impacts	Project Contribution
Stimulate new research perspectives for the humanities and social science communities, promote further the use of digital cultural heritage allowing its reinterpretation towards the development of a new shared culture in Europe.	<i>ArchAIDE</i> allows the creation of new archaeological data that will be made available both through the interactive app (supporting data access and visualization) and published as open data in a section of the ADS archaeological repository. In this way, the large amount of data produced by the <i>ArchAIDE</i> System will be available to the archaeological community as a whole, and it will help to broaden the horizons both of archaeological research (new typological studies about archaeological pottery, trade flows and their economic impact, geographical diffusion, etc.), and of all those related areas (educational, touristic, entrepreneurial, etc.) that draw on open archaeological data to produce derivatives content such as storytelling, educational tools, etc. Moreover you must consider that pottery was produced in all historical ages and in all geographic areas, therefore the data produced by <i>ArchAIDE</i> have not chronological or spatial limits and they can be used by a wide range of users
Provide innovative and creative methods for approaching cultural assets and generate applications and services to access and exploit the rich and diverse European digital cultural heritage in a sustainable way.	The project is based upon the current methodology used by archaeologists for recognising and describing potsherds, but it proposes a simplification of their work. This will allow a wide use of the <i>ArchAIDE</i> system as it does not require any additional proficiency beyond the skills already possessed by professionals or scholars. Moreover, this new approach will reduce significantly the costs of archaeological practice and, at the same time, it will greatly reduce the time of publication of the documentation, which will more quickly become part of our, both digital and material, cultural heritage. Moreover, the availability of open archaeological data always updated and enriched allow the use of the data themselves or aggregated with other digital contents, for the development of digital applications on

	cultural heritage. In fact, the data produced by <i>ArchAIDE</i> can generate an ecological data recycle, which would make the initial investment more sustainable and profitable.
Foster collaboration between those with primary expertise in the interpretation of cultural data and researchers with complementary expertise in digital and interactive frameworks.	The collaboration between archaeologists, as primary experts in the interpretation of cultural data, mathematicians and ICT scientists will be definitely strengthened by the complementarity with which these figures will be framed in the organization of the project. The proposed project is by its nature a multidisciplinary effort, originated by specific archaeological requirements and applying innovative ICT methodologies and algorithms to design a new innovative process based on mobile interface and sophisticated data processing and archival approaches. We expect that this experience, in addition to direct cooperation to achieve the goals set by the project, will foster both new interdisciplinary research, and the emergence of competitive start-ups.

2 Update of the plan for exploitation and dissemination of result

The consortium started to work on the plan for exploitation and dissemination of results. A first working document has been prepared by CNR, circulated among the partners and extended with contributions from the partners. The content is preliminary, but the report already discusses most of the topics that will be part of the final exploitation plan (such as: IPR management issues and policies; commercial exploitation, discussing the state of the art of mobile apps, possible commercialization potential for the ArchAIDE results, policies for managing the property of the outcomes, common and individual exploitation strategies). The current table of content of this document (D10.5 Exploitation Plan, to be delivered on M33) is presented in the following figure.

Among the topics described in the current draft, we resume here the following:

- *Intellectual Property Rights Management*: this topic is slightly more complex than in usual EC projects, since we have to treat both technical outcomes and data/archive-related issues. We adopt consolidated practices to define our IPR management policy related to the *technical outcomes* of the project (definition of technical results, of background and foreground knowledge, formulation of proper access rights to them, etc.). We define the rules of use and the protection measures for the project's logo and brand, since the future exploitation of the project results will require to keep and exploit also the ArchAIDE branding and visibility reached by the project. Then, we discuss and define our policy concerning the *archaeological data* used by the project tools. This encompasses both the data and knowledge contained in the *pre-existing catalogues* (which are going to be digitized by ArchAIDE and transformed in project assets) and the *data produced by the users* of our tools (i.e. data produced while documenting excavation findings). As far concern the latter, we should consider the rules for ingesting and reusing those data, since they might become another important asset for increasing the knowledge base stored in the ArchAIDE archive.

ArchAIDE D10.5 Exploitation Plan

Table of Contents

1	Abbreviations	4
2	Executive summary	5
3	Introduction	5
4	IPR Management	5
4.1	IPR Management - Technical results	6
4.2	IPR Management – ArchAIDE logo	7
4.3	IPR Management - Pre-existing Catalogues	8
4.4	IPR Management - Data produced by users.....	9
5	Commercial exploitation	10
5.1	ArchAIDE database	10
5.2	Market of mobile apps & ArchAIDE.....	10
5.2.1	Marketing opportunities for Mobile Apps.....	10
5.2.2	Mobile app for the Archaeology domain.....	13
5.3	Exploitation strategies - Whole consortium.....	17
5.3.1	Commercializing the ArchAIDE app	17
5.3.2	Commercializing or licensing derived components or apps	19
5.3.3	Policy for determining the property of the outcomes of the project.....	19
5.3.4	Active participation to the commercial exploitation activity	20
5.4	Exploitation strategies - Individual partners	20
5.4.1	UNIVERSITA' DI PISA (UNIFI).....	20
5.4.2	CONSIGLIO NAZIONALE DELLE RICERCHE (CNR).....	21
5.4.3	TEL AVIV UNIVERSITY (TAU).....	22
5.4.4	UNIVERSITY OF YORK	22
5.4.5	UNIVERSITAT DE BARCELONA	23
5.4.6	UNIVERSITAET ZU KOELN	24
5.4.7	BARAKA ARQUEOLOGOS S.L.....	24
5.4.8	Elements. Centre de gestió i difusió de Patrimoni Cultural.....	24
5.4.9	inera srl.....	24
6	Conclusions	25
7	References	25

Table of content of the current working version of D10.5 *Exploitation Plan*

- *Commercial Exploitation.* The related section of the current draft presents the preliminary ArchAIDE plan for the exploitation of the project results, defined as *foreground* having potential for industrial, commercial, or research applications. Applications include future research activities, the development, creation or marketing of a product, or processes for creating or providing services. Therefore, there are several assets which can become subject of exploitation:
 - *The ArchAIDE Database.* ADS will archive and disseminate the components of the comparative pottery collection database which could be licenced under our terms of use and access (or an appropriate CC licence), but not the database content created by users. ADS archives can be used by commercial entities, as long as the data is not re-sold.
 - *ArchAIDE and the current apps market.* We have started studying the market opportunity for mobile apps. We have drafted a review of the status of the market for generic mobile apps, which is showing an impressive growth and diffusion, both at the level of personal

apps (gaming, fitness, etc.) and professional apps. Market figures are extremely positive in this domain, with an impressive growth in sold units and revenues. The market is demonstrating a progressive transition from the usual desktop platforms to the mobile world, even for professional applications. At the same time, the status of the apps developed for the professional archaeology market is not consolidated. We have done a review of the apps developed for the professional archaeology market and we were able to review in the report only five apps (none of them having the functionalities we are designing in ArchAIDE). This can be seen as a good or a bad news: good because the competition is scarce and thus ArchAIDE could enter in a market which is still very open; bad because the lack of commercial proposals could be related to the perceived small business value of the archaeology-related market. We discuss in the report that archaeology has been so far the disciplinary context in Digital Humanities presenting the faster adoption rate of new technologies (e.g. GIS or 3D scanning). Moreover, as a professional market, it is not so small (with more than 33.000 archaeologists working in Europe). Therefore, we are inclined to see the scarce competition in the apps market as an opportunity rather than a threat.

- *Commercialization of an ArchAIDE app.* This is a first potential source of business and potential revenues for ArchAIDE. The consortium (or part of it) will evaluate the possibility to develop and commercialize the *engineered version of the prototype system* developed in the project. At the end of the project the consortium will have produced clear figures on the impact of the technology implemented in the specific application domain and on the assessment of each specific component (that will possibly signal small modifications required to make it more robust or better fitting the expectations of the users). Transforming the ArchAIDE prototype into a commercial version will require some work to ensure the required level of robustness and to reach the expected TR level. Therefore, immediately after the end of the project Inera and the other partners interested in playing an active role in the commercialization will take a decision on how to proceed with the commercial development of the ArchAIDE system.

A formal business plan will be carried out by the industrial partners and communicated to the rest of the consortium. This plan will include an accurate evaluation of the costs needed to perform the engineering of the system and the required marketing costs. Part of this plan will also be the evaluation of the costs related to the digitization of other catalogues (in the case that the experiments and assessment performed in the final phase of the ArchAIDE project will suggest the addition of other catalogues) and their inclusion in the overall system.

A price estimation of the whole system as well as the possible separate subsystems will be part of this market-related activity. We are considering two methods of purchase/business: the direct purchase of the app and a second model based on an annual fee.

Finally, aiming at a commercialization, we should determine which partners will be interested in an active participation to the future business. In the case that only INERA will be part of the business, INERA will acquire the ownership of the project outcomes from the related partners (following the policy we have defined for IPR management). INERA will proceed with the commercial activities using his own pre-existing technical and commercial staff. Conversely, in the case that other partners will be interested in taking part to the product finalization and commercialization activities, INERA and those partners will create a *spinoff* specifically oriented to the new product.

- *Commercialization or licensing derived components or apps.* The re-use of the *single results or components* is a second exploitation option. Just to give some practical

examples, this may include portions of the mobile app, the similarity-based search, the sharing and cooperation environment, and the catalogue digitalization technologies. These components will have an economic value per-se and could be used to produce new applications in the cultural heritage, didactic and tourism domains. Therefore, the results achieved by the project might be used as the basis for a tool-box to build new applications targeting visitors of museums and archaeological sites, or to contribute to the development of apps or products suggested by third parties.

- *Exploitation strategies put in place by individual partners.* Finally, this last section of the deliverable presents the individual interests and potential actions towards exploitation which will be put in place by each single ArchAIDE partner.

3 Deviations from Annex 1 and Annex 2

Explain the reasons for deviations from the DoA, the consequences and the proposed corrective actions.

3.1 Tasks

	DoA	Deviation
Task 5.3	<i>UCO, UB and UNIPI will insert the data produced in the digitization effort in the project database, following the specification defined in WP2 and WP3. UoY will train the staff to use the database and to perform data entry on the provided resources. The workflow from digitization to population has to be developed very fast so that the partners can use it in early stages of the project</i>	As we continue the work on populating the database an end of WP5 in month 20 seems not useful. With the progress of the project it has become evident that the quantity and quality of data requested by WP6 must be wider than estimated. This means that together with the digitisation of paper catalogues, whose acquisition is at an advanced stage, we need to take more pictures of potsherds. The work on the Amphorae and <i>Terra Sigillata</i> can be done quite efficiently because these pottery classes are very well structured and edited over decades of pottery research. Nevertheless, the number of pictures of different form types per class that we need is not enough. Consequently, the archaeological partners found new archaeological excavations from which new photomaterial can be obtained. However, the agreements with the Principal Investigators of these excavations taken over the last few months have led to an increase in time. On the other hand, the work on the medieval and postmedieval pottery, which is of fundamental importance for developing the appearance based similarity training in WP6, was more complex than expected. Medieval and postmedieval pottery data are less

		<p>structured and classification of the medieval pottery has to be done partially by our project. Deliverable 6.1 has not been achieved on the due date. The deviation is due to the lack of training data. We therefore were not able to train the neural network classifiers on real pottery data. More photodata of real potsherds, which will be produced over the next months, is needed for appearance based similarity training and evaluation. Also the import of data processed in WP6 has to be done till the end of the work package. We think that a constant population over the coming year with more metadata, photos and with the produced data of WP6 will increase both the quality and quantity of the data in the reference database. Therefore, we propose an extension of WP5 to month 28. The date of the deliverable of WP5 will not be affected by this extension as well as it will not affect the work on WP7 and 8.</p>
Task 6.1	<p><i>“The search engine would require the development of a suitable similarity method that will link artifacts of similar appearances. The computed similarity would be invariant to both the imaging conditions (pose, illumination, resolution) and to the preservation quality. The work will start from the scenario defined in WP2-Task2.1 specification for data representation schemes implemented in WP2-Task 2.3. Similar to work in other perceptual domains, the similarity would be learned from examples that are tagged as similar or not-similar. Deep Learning methods would be employed, and specifically Convolutional Neural Networks. The results of this Task will be reported in Deliverable D6.1.”</i></p>	<p>The main deviation we have from the original plan is due to the lack of training data. We, therefore, were not able to train neural network classifiers on real pottery data. Instead, we trained generic similarity networks on auxiliary tasks, and shifted a lot of our attention to task 6.2 (shape based matching). In a joint work with the archaeologists, we were able to reach an acceptable definition of shape similarity and sherd alignment, and this enough for us to start working on generating the data required for training the machine learning. Once there will be enough data for appearance based similarity training and evaluation, and once there will be a definition, we’ll go back to working on the appearance based matching. Since we have experience with the relevant tools, there is room for optimism regarding the time it would take to complete task 6.1.</p>
Task 9.2	<p><i>“The project logo will be trademarked at the EU Office for harmonization in the Internal Market, in order to franchise the ArchAIDE innovation results, and to permit business exploitation”</i></p>	<p>The project logo has been created and the different possibility for trademarking it have been studied. At the moment, only UNIPI is interested in trademarking it. Consequently, UNIPI, as coordinator, decided to postpone the trademarking until the applicable commercial categories will be clearer.</p>

Task 9.3	<i>“We will send online newsletter, that will be a least every four months, and will document developments, important activities, achieved results and future plans in a friendly fashion. Newsletter will be available in pdf format through the project website.”</i>	The creation of the newsletter has been postponed. During the first year it has been decided to use a faster and friendly way of communication through social media and blogs. In addition, UNIPI, UoY and INERA decided to spend time building the "Associates" page on the website, as a different method to communicate and involve the audience.
-----------------	---	--

3.2 Use of resources (not applicable for MSCA)

3.2.1 Description of use of resource for direct personnel costs for University of Pisa

Given the changes of the Annotated model Grant Agreement, the foreseen cost of an “Assegno di Ricerca”, for an amount of €49.533,29, inserted in the Amendment procedure as Subcontracting cost, has been reported as “direct personnel cost”.

3.2.2 Description of use of resource for direct personnel costs declared as unit costs

The SyGMA System does not permit to describe the use of resources for “Direct personnel costs declared as unit costs” for SME. The Spanish SMEs of ArchAIDE Consortium (BARAKA ARQUEOLOGOS SL and ELEMENT SL) have inserted “SME owner/natural person costs”.

The table explains these data distributed as person/ month per WP

BARAKA ARQUEOLOGOS SL		Total amount declared as unit cost: 6.252,78€
WP	Person/Month	Direct personnel costs declared as unit costs
1	0,25	1.142,64
2	0,36	1682,24
5	0,47	2.158,28
7	0,13	571,34
9	0,15	6980,28

ELEMENTS CENTRO DE GESTIO I DIFUSIO DE PATRIMONI CULTURAL SL		Total amount declared as unit cost: 15.631,95€
WP	Person/Month	Direct personnel costs declared as unit costs
1	0,43	1.952,01
2	2,01	9.156,99
5	0,19	888,72
9	0,46	2.110,71
10	0,33	1.523,52

3.2.3 Explanation of the total funding amount used for dissemination and communication activities

The total funding amount used for dissemination and communication activities has been insert in Part A. The amount is €98.805,52 and includes:

- Direct personnel costs declared on WP9 and WP10

- Other direct costs linked with WP9 and WP10

WP	Personnel costs	Other Direct Costs		
9	55.503,65€	8.459,68€	Travel	1.340,68€
			Other good&services	7.119,00€
10	31.738,94€	3.103,24€	Travel	3.103,24

Conclusion

This document includes an overview of the project results towards the objective of the action in line with the structure of the Annex 1 to the Grant Agreement during the first reporting period from 1st June 2016 to 31st May 2017, including a summary of exploitable results and an explanation about how they can/will be exploited.

Currently, the project has ended its first year. Work Packages 1, 3, 4, 5, 6, 7, 9, and 10 are active. The project began on 1st June 2016 with WPs 1 Management, 9 Communication/Public engagement/Innovation, 10 Archiving and Dissemination, and 2 Methodologies, specification and design. WP2 was dedicated to define the working scenario in archaeological investigation and pottery processing, the database structure, the selection of the pottery catalogues and of test bed cases. Using the ADS Roman Amphorae database as a point of departure, WP3 Creation of the Application Database has completed the database design and has begun on incorporating multilingual vocabularies of pottery characteristics. WP4 Technologies for digitisation of catalogues evaluated different solutions for OCR, in order to analyse flexibility, amount of languages supported, extendibility. The advanced OCR implementation (that aims at automatically filling the database structure) is currently under finalisation. An automatic application that is able to analyse a scanned version of the drawing has been implemented. In addition, a 3D representation of the drawing is created automatically. Once defined the workflow for populating the database, WP 5 began to digitise catalogues and photograph potsherds for training the algorithms. After a first manual population of database to see functionality and bugs, the database was automatically populated with the data from the ADS Roman Amphorae digital resource. WP6 Shape and image-based similarity search and retrieval has focused on building synthetic training data of 3D vessels as well as simulating breaking them and obtaining matching sherds. This data will be used in order to develop the algorithms necessary for matching sherds with catalogue-based drawings of vessels. While the database of ceramics was populated, an initial overview of the data was done to identify the sort of similarity to be detected. When doing so, it was discovered that there are some problems with appearance based similarity. Consequently, the maximum effort was allocated on shape-based similarity, which allows much more freedom in gaining the data. WP7 The mobile tool and Front-end Desktop Application achieved the Reference Database API for the retrieval of the ceramic types from the mobile tool through the search engine produced in WP6, started the implementation of the core components of the mobile tool, and of the "Result database" (server database supporting the mobile tool). WPs 1, 9, and 10 will last during all the project lifetime: WP1 guaranteed the correct implementation of the overall management, WP9 the communication strategies, whereas WP10 the dissemination and promotion of the results of the project.