



## D2.2: Second Report on Users' Needs

### Authors:

Hannes Selhofer, SRFG

Guntram Geser, SRFG



Ariadne is funded by the European Commission's 7th Framework Programme.



## D2.2 Second Report on Users' Needs

Version 2.1

February 2015

Versions	Nr.	Authors & changes made	Date
Draft	1.0	H. Selhofer, G. Geser (SRFG) – version for internal review	31.01.2015
Revised	2.0	H. Selhofer, G. Geser (SRFG) – revisions & corrections based on comments by Holly Wright (internal quality control)	03.02.2015
Revised – final version	2.1	Further comments/edits by H. Wright implemented, adaptations in structure of Section 4, additional evaluation of DANS-EASY	20.02.2015



ARIADNE is a project funded by the European Commission under the Community's Seventh Framework Programme, contract no. FP7-INFRASTRUCTURES-2012-1-313193. The views and opinions expressed in this presentation are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission.

## About this document

This document is a contractual deliverable of the ARIADNE research project (D2.2). The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7-INFRASTRUCTURES-2012-1) under grant agreement n° 313193.

Partner in charge of the deliverable:	Salzburg Research Forschungsgesellschaft mbH (SRFG)
Authors:	Hannes Selhofer, Guntram Geser (SRFG)
Contributions (evaluation of selected research archives and portals, identification of best practices)	<p>ADS – Archaeology Data Service, University of York</p> <p>AIAC – International Association of Classical Archaeology</p> <p>ARHEOVEST – Arheovest Association, Rumania</p> <p>ARUP-CAS – Institute of Archaeology at the Czech Academy of Sciences</p> <p>ATHENA RC - CETI – Athena Research and Innovation Centre, Cultural and Educational Technology Institute</p> <p>CYI-STARC – The Cyprus Institute</p> <p>DAI – German Archaeological Institute</p> <p>DANS-KNAW – Data Archiving and Networked Services at the Royal Netherlands Academy of Arts and Sciences</p> <p>DISCOVERY – The Discovery Programme, Ireland</p> <p>HNM NHPC – Hungarian National Museum National Heritage Protection Centre</p> <p>INCIPIT-CSIC – Institute of Heritage Sciences, Spain</p> <p>INRAP – Institut National de Recherches Archéologiques Préventives</p> <p>KNAW-DANS – Data Archiving and Networked Services at the Netherlands Academy of Arts and Sciences</p> <p>LU – The Faculty of Archaeology of the University of Leiden</p> <p>MiBAC-ICCU</p> <p>MNM-NOK – The Hungarian National Museum National Heritage Protection Centre</p> <p>NIAM-BAS – National Institute of Archaeology and Museum, Bulgarian Academy of Sciences</p> <p>OEAW – Austrian Academy of Sciences</p> <p>PIN – Pin S.c.r.l. - Servizi didattici e scientifici per l'Università di Firenze</p> <p>SND – Swedish National Data Service, University of Gothenburg</p> <p>ZRC SAZU – Institute of Archaeology at the Slovenian Academy of Sciences and Arts</p>
Quality review:	Holly Wright and Julian Richards (ADS - Archaeology Data Service)

## Table of contents

<b>Table of contents .....</b>	<b>3</b>
<b>1 Executive summary.....</b>	<b>5</b>
1.1 About this deliverable .....	5
1.2 Evaluation results .....	6
1.3 Conclusions and recommendations for the design of the ARIADNE e-Infrastructure .....	9
<b>2 Introduction .....</b>	<b>11</b>
2.1 Background and objectives .....	11
2.2 Methodology .....	12
2.2.1 The lead-user approach .....	12
2.2.2 Organisation of the work .....	13
2.2.3 Definition of key terms .....	16
2.2.4 Conceptual framework: the four level scheme of users.....	18
2.3 The baseline – main results of the ARIADNE Online User Survey .....	19
2.3.1 Main conclusions .....	19
2.3.2 Adding detail to the general picture – from D2.1 to D2.2 .....	21
<b>3 Research practices in the use of digital data .....</b>	<b>23</b>
3.1 Background information: structure of the panel .....	23
3.1.1 Professional background and experience .....	23
3.1.2 Research focus .....	24
3.2 Research practices.....	25
3.2.1 Type of data used in research.....	25
3.2.2 Main sources for obtaining research data .....	27
3.2.3 Main online sources for obtaining research data .....	29
3.2.4 Shortcomings experienced in using online archives .....	32
3.2.5 Sharing/publishing research results online .....	34
3.2.6 Other issues .....	36
<b>4 Evaluation of archives and portals.....</b>	<b>37</b>
4.1 The sample: archives and portals evaluated .....	37
4.2 General archaeology portals .....	38
4.2.1 ADS – Archaeology Data Service.....	38
4.2.2 Arachne.....	40
4.2.3 DANS EASY – The online system of the Data Archiving and Networked Services .....	45
4.2.4 Fasti Online .....	48
4.2.5 MAPPA .....	50
4.2.6 OpenContext.....	52
4.2.7 Pleiades.....	54
4.2.8 Portal to the Past .....	57
4.2.9 tDAR – The Digital Archaeological Record .....	59

4.3	Specialist archaeology portals .....	61
4.3.1	Archives Charisma – Art Database Portal .....	61
4.3.2	Bone Commons.....	63
4.3.3	ceraDAT .....	65
4.3.4	CLAROS - Cassical Art Research Online Research Services .....	68
4.3.5	Mapping Death .....	71
4.4	Portals in other domains .....	73
4.4.1	CIARD Routemap to Information Nodes and Gateways (RING).....	73
4.4.2	Eurostat .....	75
4.4.3	GBIF – The Global Biodiversity Information Facility .....	78
4.4.4	GCMD – the Global Change Master Directory .....	80
4.4.5	Morphbank: Biological Imaging .....	83
4.4.6	PhytCore phytolith database .....	85
4.4.7	Scratchpads - biodiversity online.....	87
4.4.8	SEDAC - Socioeconomic Data and Applications Center .....	89
4.5	Portals for research communication .....	91
4.5.1	Academia.edu .....	91
4.5.2	Mendeley.....	92
4.5.3	ResearchGate.....	93
<b>5</b>	<b>Lead users' suggestions for research portals.....</b>	<b>96</b>
5.1	"Ideas and suggestions" (Module C) – concept and overall results.....	96
5.2	Suggested portal approach, services and specific features .....	98
5.3	Good practice examples, experiences, opportunities for improvement .....	115
5.4	Other suggestions.....	118
<b>6</b>	<b>Conclusions and recommendations.....</b>	<b>120</b>
6.1	High-level view of the recommendations .....	120
6.2	Specific recommendations .....	121
6.3	General conclusions for the ARIADNE e-infrastructure and service development.....	124
<b>Annex I: Questionnaire &amp; guidelines provided for the portal evaluation .....</b>		<b>125</b>
<b>Annex II: Overview of portals and evaluators .....</b>		<b>135</b>

# 1 Executive summary

## 1.1 About this deliverable

### Background and objectives

This document is a deliverable (D2.2) of the ARIADNE project (“Advanced Research Infrastructure for Archaeological Dataset Networking in Europe”) which is funded under the European Community's Seventh Framework Programme. It presents results of the work carried out in Task 2.1 “Survey of users’ needs and community building”. The foundation of this document is the earlier deliverable D2.1 (First Report on Users’ Needs), which presented the results of a large international online survey among about 700 archaeological researchers and data centre managers.

The main objective of this report is to provide additional, more detailed evidence about user requirements of key target groups (“users”) of the project with regard to the ARIADNE data portal. A panel of about 25 researchers was asked to describe in detail their use of digital data archives, to evaluate existing archives and other portals and to highlight useful features of these portals which could serve as “good practices” when creating a new research data portal. This information shall support the ARIADNE project in taking informed decisions regarding the specification of the e-infrastructure and services so that they are developed in a way that corresponds to perceived and actual research needs. The mandate was to provide evidence on these issues, notably through collecting feed-back from the ARIADNE community by way of a user survey.

### Methodology

This evaluation of research portals is based on the lead user concept initially introduced by Eric van Hippel from the Massachusetts Institute of Technology (MIT) – see Section 2.2. Lead users are users of a product or service who experience specific needs earlier than the mass market and who would benefit significantly from obtaining a solution to these needs. Often, they are also early adopters of new technologies in their field. In the case of archaeological research and the use of data resources, lead users would be archaeological researchers who make intensive use of (cross-)searchable archives in their daily work, and who have therefore specific needs and a genuine interest in developing solutions to these needs.

The analysis was carried out in three phases (see Figure 1.1).

- a **preparatory phase** (drawing up a list of relevant research portals, recruiting of a panel of lead users from the network represented by the project consortium),
- **data collection** (field work), i.e. the evaluation of the selected portals by the appointed lead users – see Sections 3-5,
- and finally the **data analysis** (synopsis of the results, recommendations for ARIADNE) – see Section 6.

### Selection of portals to be surveyed

For the purpose of this survey, a portal was defined as a website that provides access to content/data of more than one organisation or project, including digital archives which curate third-party data. Most of the portals selected for the sample (14 out of 25) had an archaeological focus and provided access to content/data from several projects. Some were websites of archives and research programmes with a pre-dominant focus on one country.

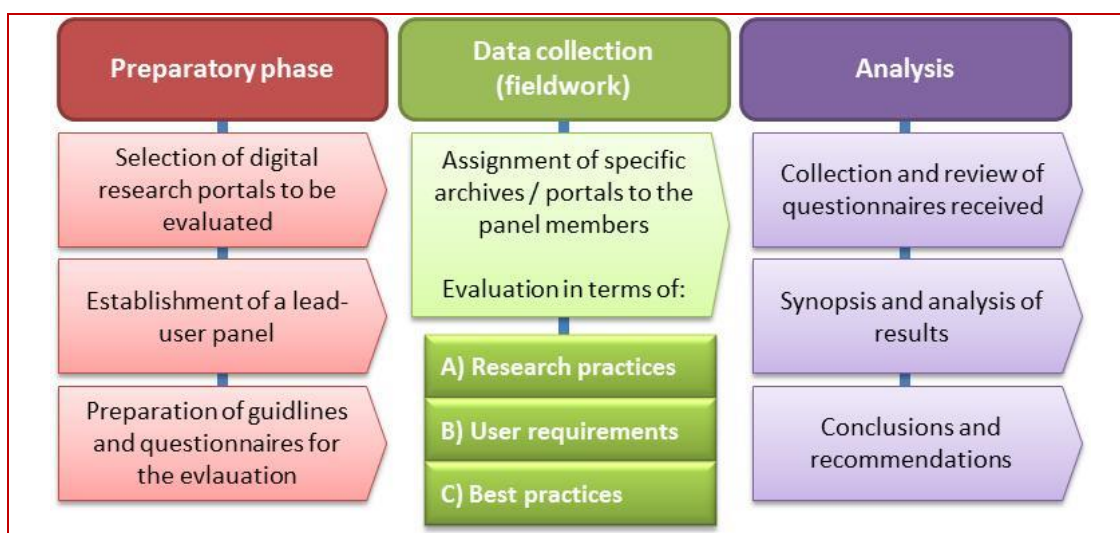
They include the best known and most advanced portals in the field (see [Section 3.1](#) for details).

The other entities were included to take account also of portals and e-research environments of other domains as well as what academic/professional networking and content sharing platforms offer to researchers. Most portals in the sample were “international”, i.e. provided access to content/data from research not only in one country.

### The panel of lead users

The lead user panel that carried out the evaluation was mainly recruited from the research organisations represented in the ARIADNE consortium. Most of the lead users are archaeologists (18 of 23); the others participants had a background in other disciplines (3) or were data managers (2). Of the archaeologists, 80% had at least 10 years of professional experience. The research focus of the archaeologists was quite diverse – both in terms of their geographic focus and research domains/specialties. About two thirds of the panel members had been using digital archives/portals for at least five years in their profession (see [Sections 3.1 and 4.2](#) for details).

Figure 1.1: Work organisation for Deliverable 2.2



## 1.2 Evaluation results

### Use of online sources and shortcomings of existing portals

We began the survey by asking lead users to describe their current research practices with regard to using online archives and databases (see [Section 3.2](#)) and to name the main shortcomings and problems they were experiencing when using them in their own work. The many responses we received to this question can be seen as a clear sign of the significant challenges that still exist in making research data available online. On the whole, the problems experienced can be grouped in the following main categories (see [Section 3.2.4](#)):

- Challenges stemming from a **lack of data quality or metadata quality** (for instance because of missing data);
- Challenges in **having access to data** (e.g. because of restricted access, technical issues how data are stored, or because of cost or copyright issues), and – as a closely related issue – **IPR issues as a barrier** for accessing/using data from online sources

- Challenges due to **problems with data integration** or with **the organisation of the archives** (e.g. difficulty in finding and extracting coherent datasets).
- **Other challenges**, including specific technical issues.

We also asked the researchers which kind of data and data sources they were mainly using and producing in their research activities, in order to validate and update the information which the Online Survey of 2013 had delivered on this issue. In fact, the answers received broadly confirm the results of the Online Survey (see [Section 3.2.1 and 3.2.2](#)). The responses demonstrate the broad variety of data needed in archaeological research. They also show how difficult it is to come up with an easy, straight forward classification for types of data, or to make a selection in what to offer, as user needs are quite diverse.

### Good practices among existing portals

We then asked the lead users to assess the strengths and weaknesses of a particular research portal and to identify good practices or innovative features that could serve as a model for the design ARIADNE portal. They provided many recommendations (see [Section 4](#) for details about the specific portals), both for specific features as well as for general principles regarding the design and functionality of research portals.

#### Good practices in the design of the portal structure

- The Archaeology Data Service (ADS) portal is seen as good practice in terms the overall structure and organisation of the portal
- Displaying further information on the right hand side of a webpage when hovering over single results or some navigation elements
- The ability to carry out a facet type browse (flexibly faceted navigation)
- Offer different routes that will eventually lead to the (same) relevant datasets, for example through searching by topic or by type of data (Eurostat portal)

#### Good practices regarding search, download and upload functions

- Semantic autocomplete in search boxes
- Display search results in different formats (images, lists, catalogues)
- Splitting the available search filters into “Content Filters” and “Technology Filters”
- An option to display the results list (of a search carried out) at one side of the web page, and a particular record on the other
- Search results should display thumbnail images of the data or representative images of the data to make it easier to distinguish the relevant data sets
- An option to search in several steps, i.e. have the possibility to refine or narrow down the search parameters after an initial, broader search has been carried out, on the basis of the (initial) search results delivered
- Image search: an image-upload-and-recognition tool (if working well) is a very useful research tool
- Tabs that guide a user to uploading and documenting the data that he/she is going to share through a portal
- Offer the user to download multiple files belonging to the same record as one coherent batch, rather than having to download individual files one by one



**Good practices regarding data organisation and presentation**

- The semantic organisation of data based on the CIDOC CRM
- The Metadata download procedure of Arachne, as download is very easy and intuitive, and resources are already formatted in XML
- Research portals should provide guidance for all potential contributors on the areas of metadata creation, licensing and standards
- The provision of content in the form of maps and coordinates
- Providing a log history for every record with its history of user inputs

**Good practices for support and added-value features**

- A “Help” section which is displayed always on the right side of the portal and guides the user through all pages.
- Allowing users to create a “MyData” or “OurData” page which aggregates standard metadata from the data shared by the researcher or research group. Various related information could be drawn in and presented based on Linked Data.
- Functions for group management such as those offered by Academia.edu and Mendeley. Group members can add literature references, post questions and comments, and receive alerts about such activities.

**Specific suggestions & learning points**

The evaluation of the portals also led to some observations and suggestions that could/should be considered in the design of the ARIADNE portal. Partly, they reflect upon weaknesses of the evaluated portals, either in terms of missing features or features that did not work properly or were not effective. These suggestions include:

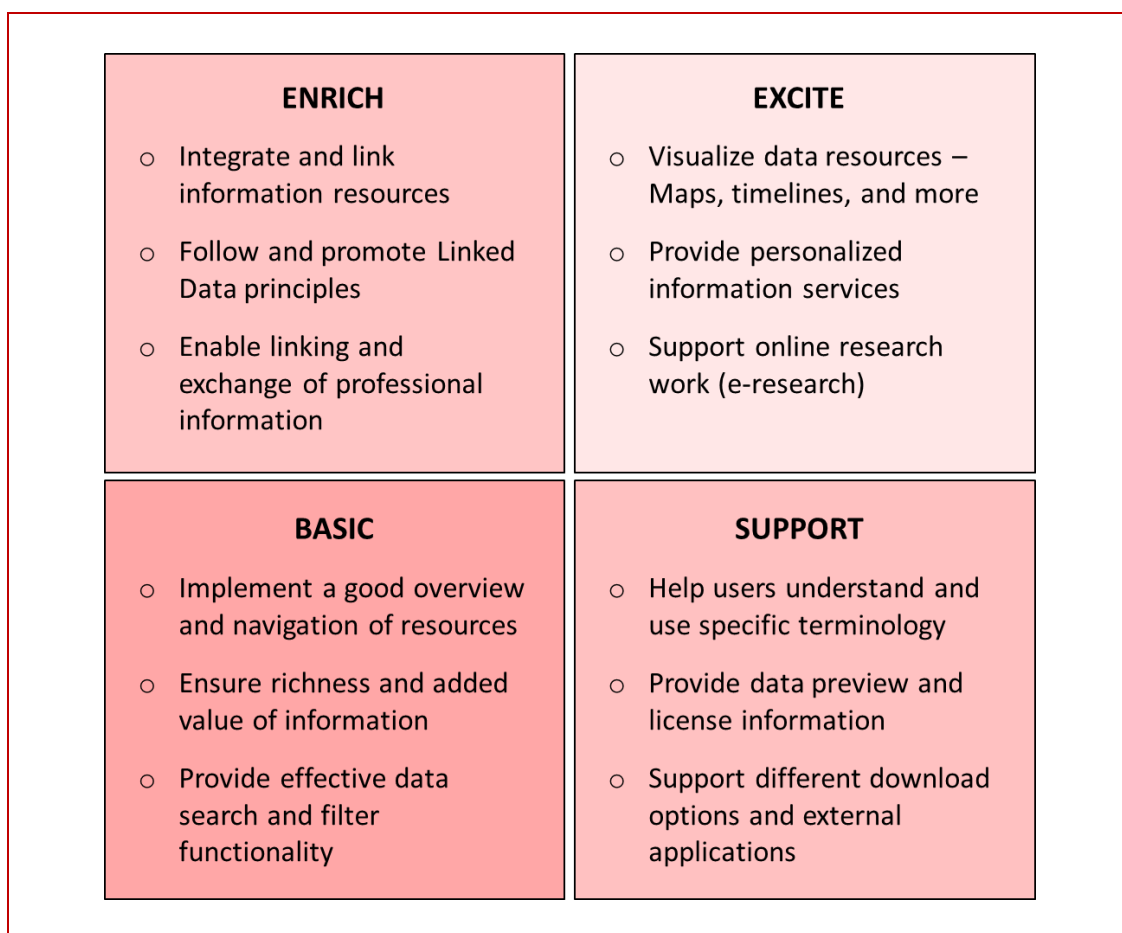
- Any portal/archive should clearly state what type of data it contains. For instance, if an archive contains only metadata, this should be clearly stated, to avoid researchers spending time searching for actual data before finding out that this is not available in the respective archive.
- Multilingual services: if different languages are offered on a portal, it should be possible to switch between languages on a page by page basis (e.g. toggle between English and German), rather than having automatic pre-set languages for different resources.
- When users complete a search, they expect to be able to download the *collated* data (e.g. so that they can see the compositional variability within that site), rather than providing a list of data for each item separately, which means that users have to copy-paste the individual data sets manually into a database.
- Enabling a structured expert exchange about objects in need of identification and thereby creating a valuable database of objects and knowledge. The same might be possible for methods and data sources.

## 1.3 Conclusions and recommendations for the design of the ARIADNE e-Infrastructure

### Recommendations for the design of research portals

Using an adapted Kano model, we have analysed and grouped the requirements of users of digital archives and portals into four categories: (i) “basic” requirements which portals must fulfil to be accepted, (ii) “support” and (iii) “enrich” features which help users to make use of the data and can thus leverage the usefulness of a portal, building upon and extending the must-have requirements; and (iv) “excite” services which are not necessarily required, but – if available – can excite broad segments of portal users (see [Sections 6.1 and 6.2](#)). In total, we identified 12 requirements (see Figure 1.2).

*Figure 1.2: Recommendations for research portals (in summary of an evaluation of portals by lead users)*



## Conclusions for the e-infrastructure

ARIADNE is the main EU FP7 Integrating Activity project in the field of archaeology. It addresses the fragmentation of archaeological datasets and limited online access to openly shared data in Europe. The project develops an e-infrastructure that will allow for interoperability of existing and newly built digital archives and, based on this interoperability, cross-archive search, access and re-use of available data. This will be a considerable step forward in the archaeological domain. The e-infrastructure will provide a common space where the currently dispersed resources can be uniformly described in the ARIADNE Registry and searched and accessed by the research community and other user groups on the Data Portal. With regard to the development of this infrastructure, the 10 conclusions can be drawn from the recommendations of the lead user survey (see [Section 6.3](#)):

1. The recommendations **confirm the overall focus** of the ARIADNE project on data discovery and access services.
2. The primary focus of the design and interfaces of the data portal should be an **overview of what data is accessible**, including statistical information on quantity, types, distribution (e.g. country/area, period).
3. The portal should focus on the **European/international dimension**. Lack of underlying resources (per country, type of data, etc.) should not be seen as a deficit, but used to promote data mobilization (e.g. implementation of national data archives).
4. Added value should also be created through **linking** data and publication **resources not held within the ARIADNE Registry** (e.g. metadata of document archives and open access publishers).
5. Linked Open Data (LOD) can play a core role for value generation, but **further uptake of LOD principles** by archaeological institutions and projects must be encouraged.
6. In the development of the data search, access and other services, **members of the user community must be thoroughly involved** and regular feedback on implemented solutions sought by the wider community.
7. **User-focused development of the portal services** and applications (relevance, usability, user-friendliness) should be at the top of the project's priorities.
8. **Services for websites for research communities** in particular subjects or geographic regions (e.g. alerts on relevant datasets) could greatly expand the reach of the data portal and, in turn, promote further data mobilization.
9. Full exploitation of the data resources (incl. metadata, conceptual knowledge) should be enabled by **interfaces for external applications** (e.g. a well-documented API, OAI-PMH target, SPARQL endpoint).
10. Support of e-research/science should, in the first instance, be provided through **integrating access to data resources** and by pointing users to existing tools for data extraction, processing and analysis.

## 2 Introduction

### 2.1 Background and objectives

#### Background

This document is the “Second Report on Users’ Needs”, a contractual deliverable (D2.2) of the EU project ARIADNE (“Advanced Research Infrastructure for Archaeological Dataset Networking in Europe”), which is funded under the European Community's Seventh Framework Programme (FP7-INFRASTRUCTURES-2012-1). The foundation of the analysis documented in this report is an earlier, more comprehensive analysis of user requirements (see deliverable D2.1, First Report on Users’ Needs), which presented the results of a large international online survey among about 700 archaeological researchers and data centre managers. This report provides complementary evidence. It focuses on good practices of existing digital research databases and archives.

The deliverable is part of **Work Package (WP) 2** “Community Building and Innovation” of the ARIADNE project. The main overall objectives of WP 2 are to help overcome the fragmentation of digital infrastructures by collecting user needs and feedback, and to foster the involvement of stakeholders in the project. The WP consists of six tasks; this deliverable is part of **Task 2.1** “Survey of users' needs and community building”, which explores current patterns of use as well as (possibly unmet) user needs of the research community of archaeologists when working with archaeological data sets. This task will be completed with the preparation and acceptance of this deliverable.

#### Objectives

The main objective of this report is to provide additional, more detailed **evidence about user requirements of key target groups** (“users”) with regard to the ARIADNE data portal. A panel of about 25 researchers was asked to describe in detail their use of digital data archives, to evaluate existing archive and other portals, and to highlight useful features of these portals which could serve as examples of “**good practice**” when creating a new research data portal. This information shall support the ARIADNE project in taking informed decisions regarding the specification of the e-infrastructure and services so that they are developed in a way that corresponds to perceived and actual research needs.

The mandate for this analysis was to provide evidence on these issues, notably through collecting feed-back from the ARIADNE community by way of a user survey. The specific objectives and goals of the analysis summarised in this document can be derived from the description of Task 2.1 in the “Description of Work” of the ARIADNE project. These are:

- to collect feedback from users on the planned data infrastructure and services (as far as its suitability to research practice and needs are concerned);
- to organise and carry out a survey on users’ needs exploring the community’s perception and reaction to the project;
- to explore (through the survey) perceived needs and expectations of researchers and the degree of their satisfaction with currently available datasets;
- to contribute to the building of a user community around the ARIADNE integrated infrastructure, in particular by making use of the survey not only as a tool to collect information, but also as a promotional tool to inform about the project.

The above listed objectives have been closely followed and addressed through the work carried out during the first and second project year in Task 2.1. The results of the various activities, in particular the results of the Online User Survey (a cornerstone of the activities during the first year), are described in Deliverable 2.1. The analysis of existing research portals by a panel of researchers (as documented in this deliverable) complements the broader survey.

## 2.2 Methodology

### 2.2.1 The lead-user approach

This evaluation of research portals for archaeologists is based on the “lead user” concept which was initially introduced by Eric van Hippel from the Massachusetts Institute of Technology (MIT).<sup>1</sup> Lead users are users of a product or service who experience specific needs months or years before the mass market will express the same needs, and who would benefit significantly from obtaining a solution to their needs. Often, lead users are at the same time early adopters of new technologies and services in their field. Translating this concept to the case of archaeological research and the use of data resources, lead users would be archaeological researchers who make intensive use of (cross-)searchable repositories in their daily work, and who have therefore specific needs and a genuine interest in developing solutions to these needs.

The lead user method is often applied in market research to assist companies in developing breakthrough products, i.e. products characterised by a significant degree of innovation. This is also the case for the ARIADNE project; if it can realise its vision of linking different research infrastructures, the innovation potential for the research community would be significant (see also D2.1 and D2.3).

*“Accurate marketing research depends on accurate user judgments regarding their needs. However, for very novel products or in product categories characterized by rapid change—such as “high technology” products—most potential users will not have the real-world experience needed to problem solve and provide accurate data to inquiring market researchers.”*

*Erich van Hippel (1986): Lead Users. A Source of novel product concepts, in: Management Science, Vol. 32, S. 791-805 (quote is taken from the abstract)*

We therefore recommended in Deliverable 2.1 that the ARIADNE project should make every possible effort to identify such lead users and work closely with them. We suggested the project should try to establish a lead user community and a platform where ideas can be exchanged with this community and thus be tested with regard to their acceptance. This would help to identify detailed, specific user needs, indicate possible solutions to these needs, and help the technical project partners to develop solutions that effectively respond to these needs. Moreover, these lead users could serve as ambassadors for the whole project and the future e-infrastructure it is going to develop.

The evaluation of existing digital research archives and portals by a group of senior archaeologists is meant as a first step in this direction. The members of the panel who have contributed to this deliverable by sharing their experiences in using digital archives, and commenting on the usefulness as well as shortcomings of existing solutions, could be a starting base to create a wider community for the discussion of related issues. We recommend that the ARIADNE project should create an online platform for discussing user needs and experiences with regard to digital archives, so that the initial work of this lead-user panel can be opened up to the community and widened. We also recommend that the Special Interest Groups should be involved in this activity.

The lead-user methodology typically involves four major steps, as suggested by van Hippel (see p. 797f., op.cit.):

1. Identifying an important market trend or technical trend;
2. identifying lead users who lead that trend in terms of (i) experience and (ii) intensity of need; analyse lead user need data;
3. project lead user data onto the general market of interest.

---

<sup>1</sup> see: Erich van Hippel (1986): Lead Users. A Source of novel product concepts, in: Management Science, Vol. 32, pp. 791-805.

We have followed these steps closely in the process designed for our analysis in the ARIADNE project, as described in the following section. The first step was a given; in the case of ARIADNE, the “important trend” is the underlying rationale of the project *per se*: the increasing use and potential of digital data archives in research. The identification of lead users, in our case, meant the selection of experienced researchers from among the research institutes involved in the project. The analysis and projection of the lead user data was carried out through a questionnaire survey and assessment of existing digital archives.

Often, the lead user concept is implemented through focus group workshops. This would have been an alternative option for the analysis at stake. Instead of carrying out one or several workshops, we asked our panel of lead users to carry out, individually, an evaluation of the existing digital research archives and portals. We went for this option as we hoped it would provide more systematic information about what types of features and solutions are seen as best practices, and where the major shortcomings are. We acknowledge, however, that there is necessarily a trade-off between collecting data through workshops (where several lead-users can discuss issues, and thus provide feed-back) and collecting data from individual lead users in parallel. While workshops would provide a better validation of individual views (through direct feed-back from other participants, and through discussion), the portal evaluation had the advantage of enabling a broader and more systematic assessment of the existing “products” in the market. We suggest that the ARIADNE project could now organise a lead user workshop (involving some of the existing panel members as well as “new” lead users from the community) on the basis of this report, to discuss the findings, validate the initial recommendations and possibly develop new suggestions for the project.

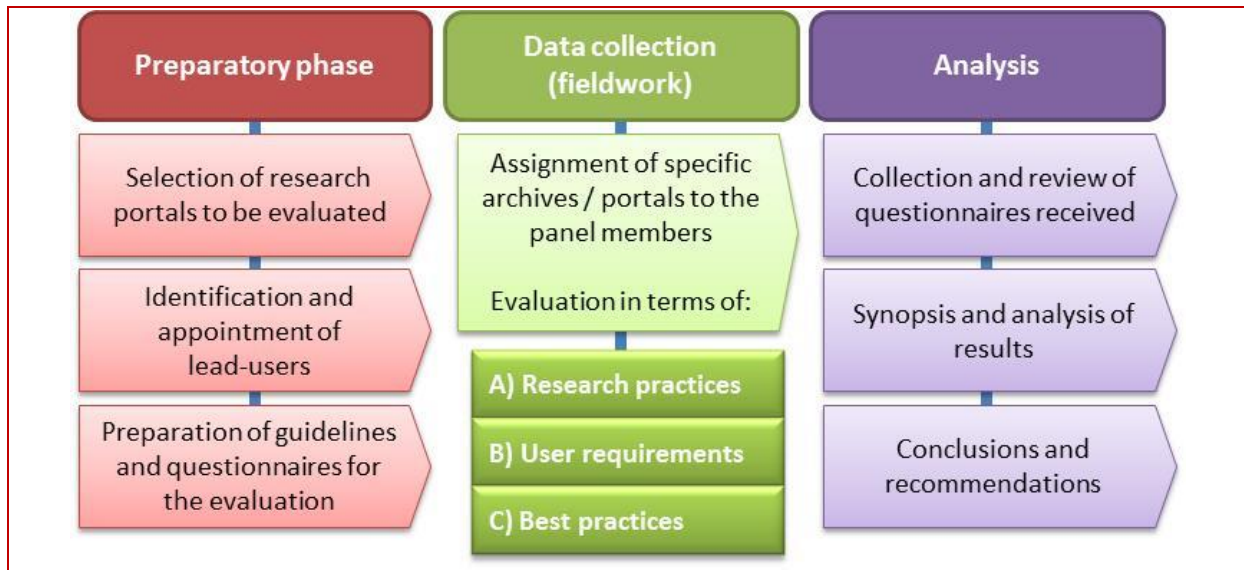
Indeed, van Hippel himself encourages researchers to adopt the proposed general methodology in different ways, depending on the context and purpose: *“I suggest that interested practitioners have no hesitation about experimenting with the general methodology describe here. (...) Researchers who wish to systematically explore the value of lead user methods will find many possible approaches.”* (p. 803)

## 2.2.2 Organisation of the work

The lead-user analysis was carried out in three phases (see Figure 2.1).

- The **preparatory phase** consisted of drawing up a list of relevant research portals as a sample for the evaluation (see Annex II), in recruiting a panel of lead users from the network represented by the project consortium, and in preparing detailed guidelines and questionnaires for the lead users on how to carry out the evaluation (see Annex I).
- The **data collection** (the field work) started with the matching of portals to be evaluated and lead users from among the panel. Subsequently, the lead users carried out the evaluation and provided the requested information about their research practices and needs with regard to using digital data. This phase was completed with the returning of the questionnaires to the central study team.
- The **data analysis** started with the collection and review of the 23 questionnaires received (see Section 3.1 for details about the structure of the sample). In some cases, the lead users were asked for clarifications or additional evidence. On this basis, an analytical synopsis of the results was prepared; the findings of this step are reported in Sections 3-5. Finally, conclusions and recommendations for the ARIADNE project were developed on this basis (see Section 6).



**Figure 2-1: Work organisation for Deliverable 2.2**

In the following, we provide some further information on the main building blocks of this analytical grid.

### **Selection of archives and portals**

The central study team from Salzburg Research drew up an initial sample of 18 digital research archives and portals which was suggested as core for the evaluation. The selection was based on existing domain knowledge and recommendations from among the consortium. We believe that this sample contains not only the most relevant portals in the field, but also a well-balanced mix of archives and portals from different domains and countries. A practical consideration was also given to the language issue, i.e. to have as many portals as possible which are operated in languages that are represented among the consortium. In short, the goal was to draw up a sample of archives and portals that has an international dimension, represents the state-of-the-art, and includes examples from different domains of archaeological research and related disciplines.

The sample consisted of 13 leading digital archives and portals for archaeologists (for instance ADS, EDNA or Fasti Online), 5 portals from related domains such as CIARD-RING (research on food & agriculture) or GBIF (the Global Biodiversity Information Facility with data about biodiversity). The list of portals that was finally evaluated is shown in Annex II.

### **Recruitment of lead users**

In parallel to selecting the main portals, Salzburg Research, with support of the consortium partners, recruited a panel of lead users according to the definition stated above (“... archaeological researchers who make intensive use of (cross-)searchable digital archives in their daily work, and who have therefore specific needs and a genuine interest in developing solutions to these needs ...”). All research partners from among the consortium were asked to identify one or several colleagues from among their institutes or wider networks who comply with the criteria for lead users, and to brief them about the work to be done.

On this basis, a lead-user panel of 23 persons was established. 18 of them are archaeological researchers, three have their professional background in a different discipline (mostly social sciences) and two panel members are data managers of research repositories. More information about the background and experience of the panel members is provided in Section 3.1.1.

## Preparation of the questionnaire guideline & instructions

The central study team then developed guidelines for the lead users on how to carry out their assignment, and a questionnaire that served, at the same time, as a template for developing their reports (see Annex I).

The questionnaire was structured into three modules.

- **Module A** explores **current research practices** and patterns of the lead users, in particular concerning their use of archaeological data. The results of this part of the survey are presented in Section 3.2.
- **Module B** focuses on the **evaluation of a specific portal**. Lead users are asked to assess strengths and weaknesses and identify good practices – the results are summarised in Section 4.
- **Module C** is about general **suggestions and ideas** how data portals could or should be improved in the future, irrespective of the specific portal evaluated in Module B. The recommendations obtained from lead users in response are presented in Section 5.

## Field work (portal evaluation)

The field work started with the assignment of specific portals to members of the lead user panel. Salzburg Research had initially suggested to each organisation (represented by specific panel members) which of the pre-selected portals their lead user(s) should assess, considering linguistic and domain specific aspects. The researchers who were assigned the task by their institute were given the opportunity to make alternate proposals, however, regarding the preferred portal they would like to evaluate. Most lead users agreed with the proposed portal, but some made use of the opportunity to suggest others.

The field work was then carried out in November 2015. The lead users (the researchers and data base managers) filled in the questionnaire, providing information about their research practices, and feed-back on the specific research portal they assessed. Salzburg Research coordinated this activity and provided help-desk support to lead-users in case they had questions.

## Analysis of the results

Most of the lead-users returned their questionnaires by the end of November; some were received in December 2015. Salzburg Research reviewed the received questionnaires and went back to the lead users on specific issues.

The results were then analysed. A synopsis of the information obtained was prepared for this deliverable (see Sections 3-5). On the basis of this synoptic view, conclusions were drawn and recommendations for the ARIADNE project were developed (see Section 6).



### 2.2.3 Definition of key terms

WP2 of the ARIADNE project is an analysis of user needs with regard to data in archaeological research. It is therefore necessary to specify our understanding of some essential terms such as “users”, “data” and “needs”. All of these terms are widely used in everyday language in different contexts; one could therefore assume that their meaning is intuitively clear. Unfortunately, terms which are used in many different contexts tend to be vague when applied to a specific context. We have described our notion of these key terms in Section 4.2 of D2.1, and suggested a framework for describing the user communities of the ARIADNE project and how they relate to each other in a systematic way. In this section, we briefly summarise the main definitions and the concept, as it is also highly relevant for the evidence presented in this document.

#### “Stakeholders”

By “stakeholders” in the ARIADNE project, we refer to projects, institutions, companies or other entities that have an interest in the project (for instance because they are affected by the project outcomes), and/or entities that may have a (positive or negative) impact on project completion. This includes internal stakeholders that are actively involved in the project (the members of the project consortium, the project sponsors) and external stakeholders. External stakeholders include, in particular but not only, the targeted users of the services that will be produced (e.g. researchers, research institutions), archaeological data repositories (e.g. if contributing to the e-infrastructure to be developed), technology providers (whose tools are needed to establish the e-infrastructure) and related initiatives pursuing similar objectives.

#### “Users”

“Users” is a central term in this report. The term is often used, in the widest sense, to describe the target communities for whom the ARIADNE project sets out to develop an e-infrastructure and services based on this infrastructure. The “users” in a project context (at least to a large extent) are the equivalent to “customers” in a business context. Users of the ARIADNE project are an important segment of the project stakeholders.

Important considerations when discussing users and their requirements are:

- **Institutional vs. individual users:** “Users” can be framed and looked at from an institutional perspective (e.g. research institutes) or from an individual perspective (researchers).
- **“Customers” vs. suppliers of data:** “using” a data centre can either mean searching and possibly downloading existing data that has been produced by others and is available at the data centre (e.g. when conducting research for a project), or it can mean depositing new data which the user has produced with his/her research project.
- **Researchers vs. repositories as users:** the ARIADNE project addresses two basic user communities: the research community that carries out research projects and generates data from these projects (researchers, institutes); and the archaeological data centres and repositories (institutional, domain-specific, international) where data can be deposited and which thus act as intermediaries for sharing data among the research community. Both are users, but the requirements and expectations they have towards the ARIADNE project can be quite different.

## Needs, requirements vs. tools and solutions to address them

The terms “needs” and “requirements” are used synonymously in this report. This is a slightly different use of the terms than in engineering and software projects, where **needs analysis** often “sits alongside requirements analysis and focuses on the human elements of the requirements”.<sup>2</sup> The broader process of **requirements analysis**, from a software engineering perspective, “encompasses those tasks that go into determining the needs or conditions to meet for a new or altered product, taking account of the possibly conflicting requirements of the various stakeholders, analysing, documenting, validating and managing software or system requirements.”<sup>3</sup> For the purpose of this report, we adopt a modified version of the engineering-focused definition of requirements analysis.<sup>4</sup> Here, user requirements (or user needs) describe those conditions, services or features (whether existent or non-existent) which the various user communities desire in order to be able to effectively and efficiently carry out their professional activities. The goal of the ARIADNE project, from this perspective, is to contribute to better fulfilling user requirements in archaeological research, for instance by enabling new services and tools which meet user requirements in a better way than the existing services and tools.

We also recommend not confusing *needs* (or *requirements*) and possible *tools* (i.e. *solutions*) to address these needs. While we should think of a need as a “job to get done”, the solution is the means to carry out the job.

### “Data”

The most challenging and potentially conflicting definition is what we understand as “data” (or, rather, what is *not* considered as data in the strict sense). While this is a decision which is ultimately beyond the mandate of WP 2 and thus a bit outside the scope of this deliverable, the issue inevitably comes up during the evaluation of data archives. The challenge arises, to a large extent, from the wide use of the term “data” in many different contexts, including technical terminology as well as everyday language. This can lead to considerable confusion as to what we mean by “archaeological research data” (also in interviews with researchers). The following basic distinctions between different types of data can be helpful in this context:

- **Raw data vs. processed data.** A major distinction is to be made between raw data (i.e. unprocessed data as collected, for example, in field surveys, in labs or through experiments) and processed data (for instance aggregations of raw survey data).
- **Metadata vs. the ‘actual’ data.** Another important distinction is whether it is “data about data” (metadata) describing the structure and/or content of data, or whether the discussion is about the actual research data themselves. Both data and metadata are highly relevant for the ARIADNE project and its users.
- **Different types of data.** In archaeology, “data” can mean a lot of different things – including images, texts of different types, GIS and other location-based data, maps or technical data from lab tests.

---

<sup>2</sup> see “Needs analysis” on Wikipedia: [http://en.wikipedia.org/wiki/Needs\\_analysis](http://en.wikipedia.org/wiki/Needs_analysis) (accessed in Feb. 2014)

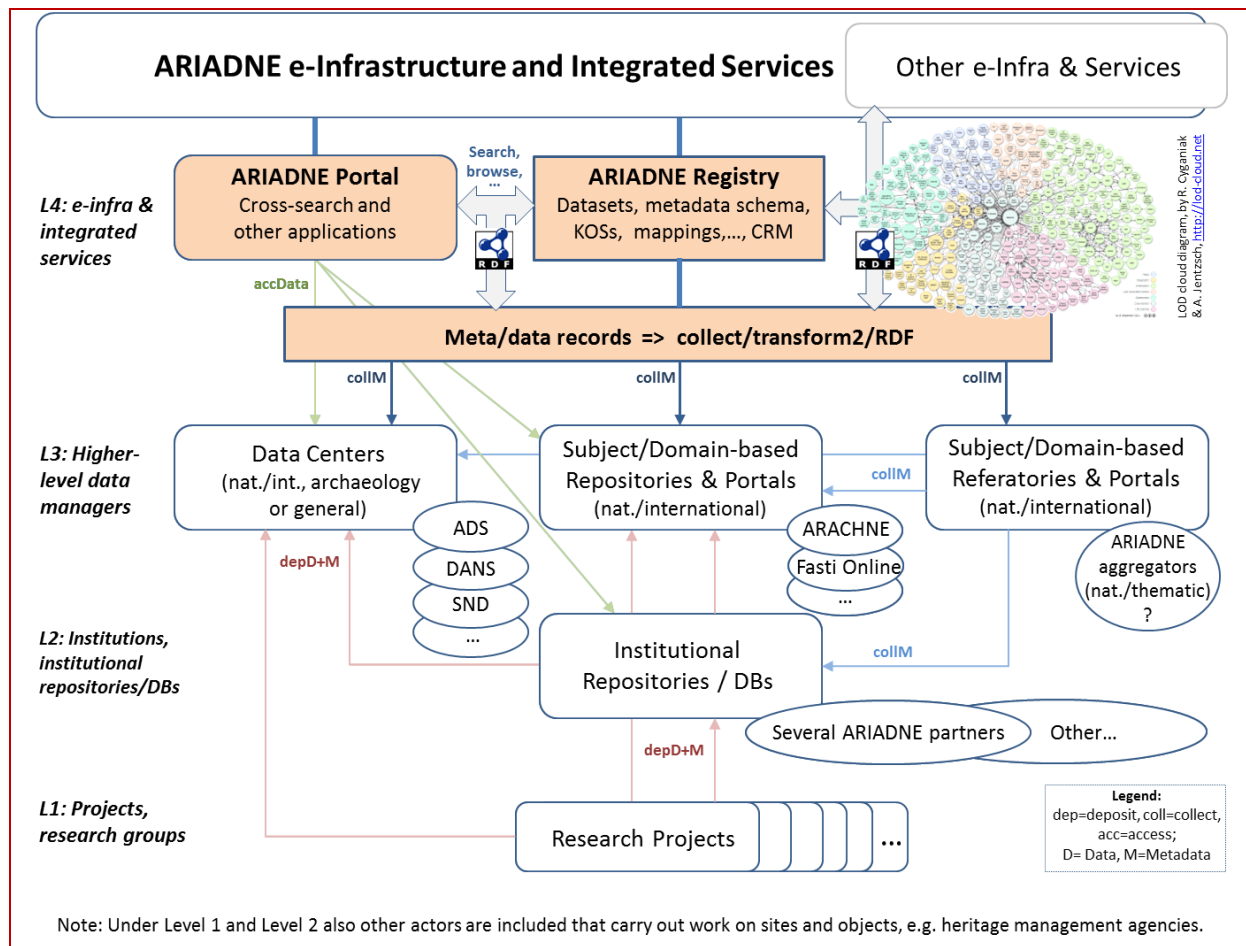
<sup>3</sup> see “Requirements analysis” on Wikipedia: [http://en.wikipedia.org/wiki/Requirements\\_analysis](http://en.wikipedia.org/wiki/Requirements_analysis) (accessed in Feb. 2014).

<sup>4</sup> for the technical work packages of ARIADNE, and in particular for WPs 12 and 13 (which also identify user requirements, but specifically with regard to the technological concepts) it may be more useful to stick to the classical concept of requirements analysis.

#### 2.2.4 Conceptual framework: the four level scheme of users

For dealing with stakeholders and analysing user requirements, we have proposed a four-level user framework in D2.1. This distinguishes four institutional levels of relevant user communities. Figure 2.2 depicts the major elements of this framework which are explained and described in the following sections.

**Figure 2-2: Conceptual framework for the analysis of ARIADNE stakeholders and users**



This framework reflects a data workflow perspective where:

1. Data is produced and managed by research projects or in the context of other work carried out on archaeological sites and objects, e.g. heritage management (Level 1),
2. the content and data or datasets together with metadata is deposited in institutional repositories (Level 2) or data centres or subject- and domain-based repositories (Level 3),
3. and the metadata from several repositories is collected (e.g. harvested) into a common metadata pool, and search and other services are provided based on the metadata (Level 3). Furthermore on this level there can also be special services that support some workflows, for example controlled vocabulary services.

For a detailed description of the four levels, we refer to Section 4 of D2.1.

The evidence presented in this report focuses on the user of requirements of **individual researchers** (Level 1) and, to a minor degree, of **managers of data repositories** (Level 2). The panel of lead users who were asked to evaluate existing research archives and portals consisted of representatives from these two communities.

## 2.3 The baseline – main results of the ARIADNE Online User Survey

This section summarises the main results of the ARIADNE Online User Survey from November 2013 (as presented in Deliverable 2.1). The findings of this survey constitute the empirical baseline and starting point for the portal evaluation. The ARIADNE Online Survey addressed the international archaeological research community, including both researchers and directors/managers of archaeological data repositories. These target groups were addressed with two different questionnaires. The roll-out of the survey to the dispersed population was carried out with support of the ARIADNE consortium, predominantly by leveraging the networks and communities of the various research partners. The final, cleaned net sample consisted of 692 questionnaires (640 researchers and research directors, and 52 repository managers).

### 2.3.1 Main conclusions

#### Overall conclusions

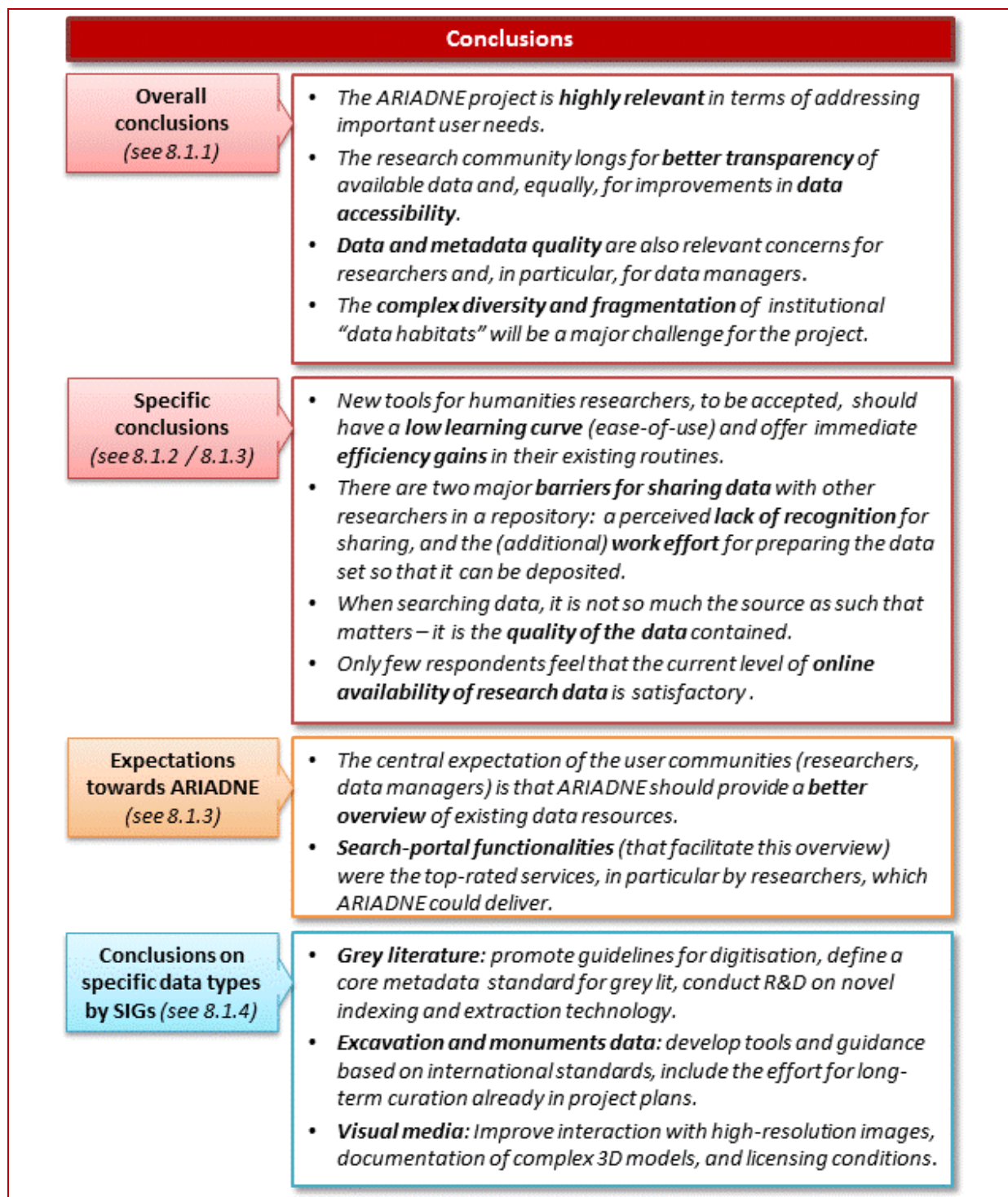
The survey results clearly confirmed the **high relevance of the ARIADNE project**, as it addresses important user needs with regard to research data which are not well catered for by existing services. More than 60% of the researchers surveyed said they were not or less satisfied with the current situation with regard to major parameters. In particular, they criticised a lack of transparency of available data, and difficulties in having access to data. Any improvements in these areas would be highly appreciated by the user community.

At the same time, the results also confirmed the **significant challenge** with which the ARIADNE project is confronted in its mission to create an e-infrastructure with services that respond to these needs. The responses documented the enormous degree of fragmentation with regard to potentially relevant data for integration, presented by a complex diversity of institutional data habitats and different types of “repositories”. To link these project archives with a common repository will require new workflows (and possibly dedicated staff) which may not be available in many research institutions.

#### Conclusions on user requirements

The central conclusions drawn from the survey regarding user requirements were:

- The research community expressed, in particular, a need for an improved **transparency** of available research data (it is difficult to know which data actually exists, due to the enormous fragmentation of data resources in the field) and improvements in data **accessibility**.
- The major barriers with regard to accessibility were **costs** (e.g. for obtaining licences to use pictures, or for subscription fees) and the problem that relevant literature and **data is often kept in other places** than where it is supposed to be (e.g. in private collections of other researchers).
- **Data and metadata quality** were further concerns of researchers; any improvements in these fields would be highly welcome.

**Figure 2-3: Main conclusions drawn from the ARIADNE Online User Survey**

Source: ARIADNE project, D2.1

### Implications for the ARIADNE project

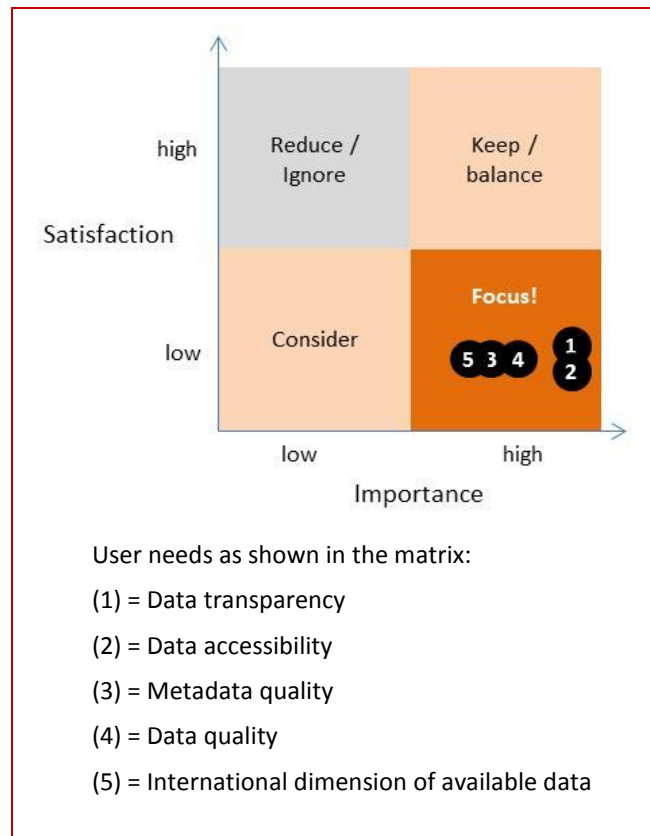
From these findings, the study team concluded that **ARIADNE has a broad field of opportunities to create real value for users**. While it is clear that the project cannot solve all problems, ARIADNE has a high impact potential if its services can deliver improvement in any of the above mentioned areas (see matrix in Figure 2.4 – all five domains of user requirements are in the segment which suggests focusing on them).



**Figure 2-4: User requirements of archaeological researchers according to their importance and the satisfaction with the existing situation in a strategy matrix**

The survey explored the perceived importance and the satisfaction of researchers with regard to five parameters for using digital data archives for research purposes: data transparency, data accessibility, metadata quality, data quality, and the international dimension of data.

All five dimensions were found to be very important, with data transparency and data accessibility ranking on top. The researcher community is not satisfied, however, with the current conditions with regard to these parameters. There is much room for improvement with regard to the availability of digital research data.



### 2.3.2 Adding detail to the general picture – from D2.1 to D2.2

While these findings were a clear confirmation of the relevance of the ARIADNE project, and a thorough empirical baseline to establish a general picture of the situation, the results from this standardised survey were not specific enough to take a strategic decision on priority areas, and to facilitate the choice and design of technical solutions, a further analysis of specific user requirements in the five domains is required. The study team therefore concluded that it needed a second analytical step where user requirements will be broken down further and analysed in more detail.

When asked about their expectations towards the ARIADNE project, many researchers expressed a hope that the resulting services can improve the transparency of what is available, the search capability and, possibly, the conditions of access (e.g. promote open access repositories). More specifically, one of the main suggestions made by many respondents was that ARIADNE should **establish a new portal for data search**. If such a new portal (on top of existing data resources) is established, users will clearly expect an added-value – i.e. it must have other and better features, or provide access to more resources. While an improved overview, cross-searching and filtering of data resources would be quite some progress on the current situation, the specific requirements were not fully clear.

This is the starting point and mission for this report which should be regarded as complementary evidence to D2.1 rather than a mere “update”. When exploring different possibilities with regard to how to analyse specific user requirements in more detail, notably with regard to the functionalities of a future ARIADNE portal, the study team finally opted to carry out an evaluation of existing research portals. The idea was to systematically collect “good practice” examples from a range of different archives and portals which could then serve as inspiration for the ARIADNE services.

This evaluation was to be carried out by experienced users of digital archives. As part of the evaluation process, the evaluators were also asked to describe (in a qualitative way) their own research activities and exactly how they make use of digital archives in this process. This provided further insights, in particular from a process perspective that helps to understand user requirements.

In the following chapters, we present the results of this portal evaluation by lead users.

- **Section 3** focuses on **current practices in using digital data** archives: how do researchers specifically make use of digital data (e.g. in which stages of a research project, which are their preferred sources, what kind of data do they use and produce)?
- **Section 4** presents the **results of the portal evaluation**: what are the strengths and weaknesses of existing portals? What are interesting features which researchers like and could possibly be considered by a future ARIADNE portal?
- **Section 5** presents a **“wish list”** of researchers and recommendations with regard to digital data archives: which types of features are most important for them, where would they, ideally, like to see improvements in the future?

### 3 Research practices in the use of digital data

The first part of the questionnaire (Module A, see Annex I) addressed the research practices of the panel members with regard to digital data. The researchers were asked to provide information about:

- their professional experience in archaeology (in general) and their research focus;
- their experience in the use of online databases or portals;
- the main sources from which they normally obtain the research data they need and, more specifically, the main online sources they are using for this purpose (e.g. specific archives);
- the main shortcomings and problems they have experienced in the use of online archives and databases in their own work;
- the typical work flow in an archaeological research project with regard to the use of online archives and portals;
- to what extent and in what way data which their research group is producing is typically being published, and to what extent data is made available from online archives.

This chapter summarises the responses we have obtained.

#### 3.1 Background information: structure of the panel

##### 3.1.1 Professional background and experience

The panel that conducted the evaluation of the digital research archives consisted of 23 persons; 18 of them are archaeological researchers, three have their professional background in a different discipline (mostly social sciences) and two persons were data managers of research repositories. The latter had a background in archaeology, but said they had no or only few of their own research activities in recent years. Therefore, they answered the questionnaire from the perspective of a provider of a research database rather than a user. As some panel members (3) evaluated more than one portal, their number differs from the number of portals that have been looked at and assessed (26 in total, see Annex II).

**Table 3-1: Structure of the lead user panel in terms of their main discipline/activity**

Professional background	Number
Archaeological researchers	18
Researchers from other disciplines	3
Research database managers	2
<b>Total</b>	<b>23</b>
Total no. of questionnaires received from them	26

When looking at the 18 archaeological researchers among the respondents, the majority of them have significant professional experience in their field, and also in the use of digital archives and databases (see Table 3.2). The group therefore is clearly qualified for the task to which it was assigned. Almost 80% of the archaeologists who participated in the evaluation have at least 10 years of professional experience, and more than 70% of them have been using digital archives and portals in their research for at least 5 years, 50% for 10 years or longer. Only one of the respondents (a senior researcher with more than 10 years professional experience) said that he/she had no experience in the use of digital data archives. The responses of this person to the questions in Module A (as reported in this section) have therefore not been considered. The responses of the



other panel members (database managers, researchers from other disciplines) have been used as far as relevant for this section.

**Table 3-2: Structure of the lead user panel in terms of professional experience**

Professional experience ...	... in archaeological research			... in the use of digital archives/portals	
	No.	%		No.	%
no experience	-	-		1	6%
1-4 years	2	11%		4	22%
5-9 years	2	11%		4	22%
10-19 years	7	39%		8	44%
20+ years	7	39%		1	6%
<b>Total</b>	<b>18</b>	<b>100%</b>		<b>18</b>	<b>100%</b>

### 3.1.2 Research focus

The research focus of the 18 archaeologists is quite diverse – both in terms of their geographic, thematic and domain focus. It is therefore unlikely that the evaluation results could suffer from a regional or domain specific perspective. Specialisations include traditional fields (classical archaeology, excavations) as well as more recent ICT-related / digital aspects (use of 3D, “theory of digital archaeology”, “applications of ICT in the field of culture”).

The following definitions of research profiles (quoted as stated by them in the questionnaires, with minor linguistic adjustments only) are indicative of the professional background of the researchers who conducted the assessment of the portals.

- *“Landscape archaeology, GIS, agent based modelling, network analysis, bronze age, iron age, high and late middle ages, northern Europe, Mediterranean.”*
- *“The main areas of my research interest include classical archaeology, excavation, ancient Greek pottery, applications of ICT in the field of Culture (databases, 3D, GIS, info kiosks, handheld guides, educational applications, etc. within museums, archaeological sites or monuments.”*
- *“Computational applications for archaeological questions; currently I am working with Aegean seals and pattern recognition algorithms. I am involved in writing best practice guides and in the planning for an online archive.”*
- *“Urban archaeology and town planning; administrative and scientific information digital archiving; geophysical/geological surveys, excavations; digital archiving of archaeological field data, findings, archaeological complexes and monuments; data bank of law-constraints; topographical lexicon of Rome; archaeological ontologies; archaeological e-communities; methodological innovation; open access & digital libraries.”*
- *“Use of 3D as a research environment in archaeology, ontology engineering, digital libraries, knowledge repositories, formal representation of reasoning, theory of digital archaeology.”*
- *“Ceramics. My research focuses on the utilization of interdisciplinary analysis in archaeological interpretation and the combination of interdisciplinary research with archaeological theory. I analyse ceramics from the Neolithic to the Late Middle Ages mainly from Hungary but I also analysed ceramics from Austria, Spain and Italy.”*

- *“Applications of ICT in the field of culture: I have been collaborating for several years with the [...] in the context of research and development projects aiming to create technology applications (info kiosks, handheld guides, educational applications, databases etc.) within museums, archaeological sites or monuments etc.*
- *“Remote-sensing, funerary archaeology, iron archaeology in Romania“*
- *“Thracian Archaeology, Balkan prehistory, Ancient agricultural technologies, Ancient economies, Cooking and food preparation in archaeology“*
- *“I am specifically interested in the sharing of geospatial data within the cultural heritage domain utilising Spatial Data Infrastructure (SDI). Also, I am currently evaluating a national strategy on the archiving and reuse of digital archaeology data.”*
- *“Archaeological archiving; semantic interoperability in Irish archaeology.”*
- *“The focus of my research institute is mostly related to the definition and application of new technologies to cultural heritage (from data acquisition to data management and publication). Concerning data management, my institute is interested in the design of digital repositories for the management of digital objects including 3D, the development of ontologies for the management of cultural heritage, the development of cross-walk mappings between metadata schemas and standards used for the documentation of cultural heritage, and the use of tools for data mapping, conversion and encoding in RDF.”*
- *“Early Medieval archaeology and history, Slavic mythology and early medieval spatial structures, computer sciences in archaeology, experimental archaeology, ethnoarchaeology, archaeoastronomy“*
- *“Archaeology of the longue durée in Italy and North Africa, the Berbers, Survey and Excavation Methodology, Slaving states, Roman slavery“*
- *“My own specialisation is intra site spatial analysis of Stone Age settlements in the Netherlands and computer applications in general (ranging from databases, GIS, statistics to long-term archiving). This involves participating in a team of researchers in the analysis and interpretation of excavations.”*
- *“Aerial and satellite images (processing, evaluation, interpretation). Further, the study of historic agricultural landscape, and historical maps.”*

## 3.2 Research practices

We then asked the researchers to provide some information about the current research in terms of data and data sources used, shortcomings experienced in the use of online sources, and about typical work flows (i.e. specifically in which stage of a research project they make which use of data). We summarise the main findings in this section.

### 3.2.1 Type of data used in research

The first part (about data and data sources) explored an issue that had already been covered by the Online Survey of 2013. While the survey asked respondents predominantly to select from predefined items (in order to facilitate a quantitative assessment), this time the respondents were given the opportunity to describe the data they are using and the sources in free text format. Even so, the answers received broadly confirm the results of the Online Survey.

In the ARIADNE Online Survey (2013), the respondents were asked to rate the importance of different types of data for their research. They were presented with a list of more than 10 types of data, including (for example) excavation data, GIS data, prospection & field survey data, and data for corpus studies (see Section 6.2.2 of D2.1). The survey concluded that the single most important type of data (if measured by the number of researchers for whom they are important) is excavation data. Seventy-five percent of the respondents said that excavation data was “very important” for them to carry out their research projects. Also very important for a large group of researchers (about 50% each) were GIS data, data stemming from material or biological analysis, and data from field surveys. The other types of data are not irrelevant; quite the contrary, they all have their users; in most cases, at least 50% of the respondents said that they were at least “rather important”

We asked the researchers in the lead user panel a similar question (which kind of data they were mainly using in their research), but without giving them predefined options. The following answers given confirm the results of the online survey and demonstrate the broad variety of data needed in archaeological research. They also demonstrate how difficult it is to come up with an easy, straight forward classification for types of data, or to make a selection in what to offer.

Excavation data clearly plays a very important role, it is mentioned in many of the answers:

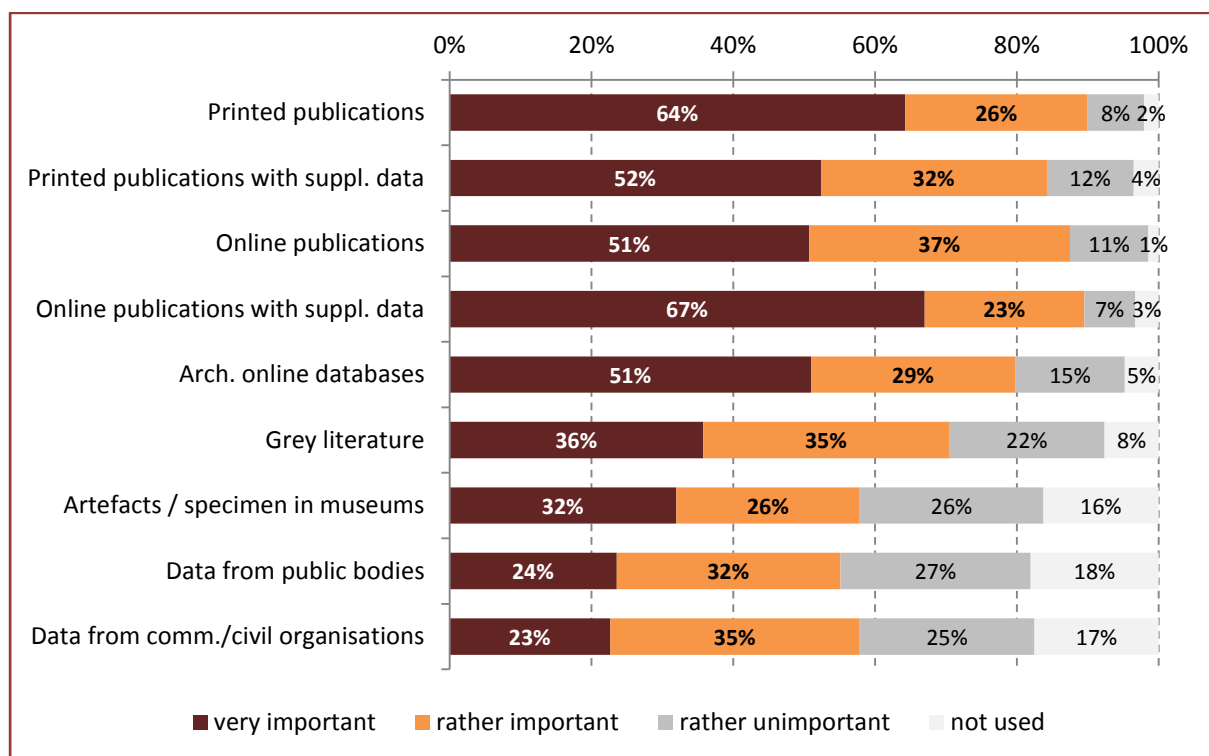
- *“Excavation data, remote sensing data, ceramic databases”*
- *“Excavation and field survey reports, literary and iconographic sources, ethnographic and laboratory data”*
- *“Remote sensing data (lidar, geophysics, aerial images), monument databases, artefactual data collection”*
- *“Remote sensing, excavation metadata”*
- *“Excavation data, field survey data, legacy data, 3D resources, GIS datasets, semantic data and metadata”*
- *“excavation data, field survey data, museum collections, folk narratives, written sources”*
- *“Excavation and survey data, texts, anthropological literature”*
- *“Normally we produce our own data within research projects. If available we retrieve existing data from previous studies – mostly field survey data and excavation - to enrich our datasets with it.”*
- *“With reference to applications of ICT in culture, I am using all kind of data a) excavation data in various formats: texts, images, 3d data, reports etc., b) museum data from artefact catalogues, c) archaeometry data, d) 3d representations, e) multimedia etc. trying to incorporate all these in ICT applications for different groups and purposes. I am also searching scientific publications or bibliographic lists with reference to archaeology.”*
- *“object description data from a database”*
- *“All type of digital resources and mainly bibliographic, archival, and catalogue data; digital collections belonging to museums, libraries, and archives, archaeological sites; 3D models and replicas; audio and video resources.”*
- *“Legacy data: Texts, images, 3D, GIS locations, scientific documentations, such as: survey reports, archaeological field data, excavation data, etc.; cartographic, graphic and photographic documentations, both historical and new documents; multimedia contents; 3D models”*
- *“Digital public data: public map-bases (vector, raster); on line literature and open access repository; open data.”*

- „3D documentation data, digital images, digital maps, 3D models“
- “I use excavation data, data from processing the finds (archaeologists' description and evaluation of ceramics) and laboratory data.”
- “The analysis of Stone Age sites does involve all of the different types of data, although for the spatial analyses of an excavation, the field data are more common.”
- “Aerial photographs, satellite photographs, maps, map portals, excavation data, field survey data”

### 3.2.2 Main sources for obtaining research data

We then asked the researchers which sources they used to obtain their data, we can once again compare the results with those of the Online Survey on the same issue. The main conclusion from the Online Survey was that *there is no single most important source* for research data. Researchers in the digital era need to be flexible and make use of all kinds of data sources, depending on where the required data are available (see Figure 3.1). It is not the source as such that matters – it is the quality of the data contained. This was also previously confirmed by the comments received from respondents (e.g. “All sources are important and must be first-hand”).

**Figure 3-1: Results of the ARIADNE Online Survey (2013) on data sources: “When working on research projects and searching for data: how important are the following sources for you and your research group for collecting data?”**



N = 543-579 (depending on number of respondents without answer)

The responses obtained for this update of the survey (from researchers among the lead user panel) confirm that all of the above sources are relevant. The answers also show that some researchers, when asked about sources, think more in institutional terms, i.e. from which organisations the data are available (e.g. “my own institute”, “cultural institutions”, “project partners” or a specific database), while other think more about different ways in which data can be archived, irrespective of the institutional context (e.g. “databases” in general, “published articles”, “online publications”). This

might be attributable to the way the question was posed, but could also indicate different research approaches. In some research contexts, the search for existing data might focus on specific sources which are known or expected to have relevant data. In other contexts, researchers are much less focused on specific organisations in the way they conduct their background research.

Answers received include the following:

- *“Preferably resources available from researchers who work in similar fields and on similar research questions (e.g. published GIS studies), but in most cases one has to rely on site location in a webGIS and add further information manually.”*
- *“My sources for research data are:
 
  - *Online or printed reports and articles*
  - *Institutional (mainly museum) databases*
  - *Project partners*
  - *Educational or cultural organisations*
  - *Museum (or other) websites*
  - *Laboratories (archaeometrical, digitisation etc.) of our own research centre”**
- *“databases, some institutional, some widely accessible”*
- *“cultural Institutions and research centres both at National and International level”*
- *“Primarily, we find our data from SSBAR archives and other legacy data of our Institute, along with all the other informative sources available on the specific matter of Public Archaeology at Rome”*
- *“our own repository and laboratory”*
- *“our own and external laboratory data”*
- *“Own data – published articles”*
- *“Institutional databases, published materials”*
- *“Government monument databases”*
- *“Excavations databases”*
- *“Data created within The Discovery Programme: databases, remote sensing. Data created through licensed archaeological activity and curated by government departments and bodies and other organisations.”*
- *“Records of governmental organisations, institutional databases, own data collected in field activities.”*
- *“own institute databases; bibliography and full text databases; written sources databases”*
- *“Print and online publications, bibliographical sites.”*
- *“The main sources of excavation data are made available by the organisation responsible for the excavation, almost always through direct personal contacts or incidentally through the archaeological data archive at DANS.”*
- *“Institutional databases, own databases”*

### 3.2.3 Main online sources for obtaining research data

We then asked the researchers to name the 3-5 main online sources from which they obtained research data. This question links to an item in the ARIADNE Online Survey of 2013, where the respondents were asked to rate the online accessibility of data. Only a few respondents felt that the online availability of research data was fully satisfactory. For any type of data only a minority of respondents (typically 5-10%) said that the accessibility was “very good”, and less than 50% said that access was at least “good”.

However, this does not mean that researchers do not make use of online sources. The respondents named numerous online sources, including in particular the portals that are evaluated in this document. The following table provides the list of all digital archives, portals and repositories that were quoted by the respondents.

The results demonstrate the importance of national initiatives. In many cases, the researchers named mostly specific national digital archives which are maintained by institutions (research organisations, government bodies) of the country in which they do their research. This is another proof of the fragmentation of data sources across different countries and institutions, which constitutes, at the same time, the main motivation for launching the ARIADNE e-infrastructure initiative, but also a critical challenge. The ARIADNE project has to take great care in developing a realistic roll-out plan for its services: which are the main existing data sources and collections to start with (i.e. which will be linked by the e-infrastructure first), and how can ARIADNE then go about adding further collections from other sources so that the e-infrastructure can grow over time?

**Table 3-3: Digital archives and repositories used by the researchers**

Name of the archive/repository	Data obtained
sardegnaeopoitale.it	Geographical data provided by Regione Sardegna
CGMA-MAGIS	GIS database on regional survey projects in the Mediterranean
EBIDAT	Database of the European Castles Institute
ZENON	OPAC data of DAI
opencontext.org	Archaeological research data base
Odysseus, Hellenic Ministry of Culture <a href="http://odysseus.culture.gr/index_en.html">http://odysseus.culture.gr/index_en.html</a>	Images of objects, sites, monuments, plans, relevant texts and bibliography
Perseus Digital Library <a href="http://www.perseus.tufts.edu/hopper/">http://www.perseus.tufts.edu/hopper/</a>	Images of objects, sites, monuments, plans, relevant texts and bibliography
Academia <a href="https://www.academia.edu/">https://www.academia.edu/</a>	Scientific publications, reports, articles
Arachne <a href="http://arachne.uni-koeln.de/drupal/">http://arachne.uni-koeln.de/drupal/</a>	XML files with object metadata
OPAC.SBN ( <a href="http://www.sbn.it/">http://www.sbn.it/</a> )	Bibliography references
Fasti Online ( <a href="http://www.fastionline.org/">http://www.fastionline.org/</a> )	Archaeological “folder” referring to field excavations and researches; excavation reports
FotoSAR ( <a href="http://www.fotosar.it/">http://www.fotosar.it/</a> )	Archaeological photos of Roman National Museum collections
Academia ( <a href="http://www.academia.edu">www.academia.edu</a> )	Archaeological/scientific data available in articles
ResearchGate ( <a href="https://www.researchgate.net/">https://www.researchgate.net/</a> )	Archaeological/scientific data available in articles
Odysseus, Hellenic Ministry of Culture <a href="http://odysseus.culture.gr/index_en.html">http://odysseus.culture.gr/index_en.html</a>	Images of objects, sites, monuments, plans, relevant texts and bibliography
Perseus Digital Library <a href="http://www.perseus.tufts.edu/hopper/">http://www.perseus.tufts.edu/hopper/</a>	Images of objects, sites, monuments, plans, relevant texts and bibliography
Academia <a href="https://www.academia.edu/">https://www.academia.edu/</a>	Scientific publications, reports, articles
CIMEC (RO) <a href="http://www.cimec.ro/Arheologie.html">http://www.cimec.ro/Arheologie.html</a>	Sites types/location

Preistorie.ro <a href="http://arheologie.ulbsibiu.ro/publicatii/bibliotheca/banat/cuprins1.htm">http://arheologie.ulbsibiu.ro/publicatii/bibliotheca/banat/cuprins1.htm</a>	Sites types/location
Enciclopedia Romaniei <a href="http://enciclopediaromaniei.ro/wiki/Index:Localitate%83%C5%A3i_din_Dacia">http://enciclopediaromaniei.ro/wiki/Index:Localitate%83%C5%A3i_din_Dacia</a>	An index of places about the Roman province of Dacia, compiled from territorial criteria and accounts for toponyms settlements
Archaeological Map of Bulgaria <a href="http://www.naim-bas.com/akb/">http://www.naim-bas.com/akb/</a>	Localization, chronology and character of archaeological sites in a given area
Corpus Vasorum Antiquorum <a href="http://www.cvaonline.org/cva/browse.htm">http://www.cvaonline.org/cva/browse.htm</a>	Iconographic sources
ADS <a href="http://archaeologydataservice.ac.uk/">http://archaeologydataservice.ac.uk/</a>	Varying range of data form survey data to grey literature reports
Geological Survey of Ireland <a href="http://gsi.ie/">http://gsi.ie/</a>	Soil and geological mapping services
Heritage Maps - The Heritage Council (IE) <a href="http://heritagemaps.biodiversityireland.ie/#/Map">http://heritagemaps.biodiversityireland.ie/#/Map</a>	Heritage viewer, range of data from under water archaeology to monuments and landscape data. Finds locations, burials, road development-led excavations
National Monuments Service Archaeological Survey of Ireland web GIS <a href="http://webgis.archaeology.ie/NationalMonuments/FlexViewer/">http://webgis.archaeology.ie/NationalMonuments/FlexViewer/</a>	Locations, place names, monument reference identifiers, monument classifications, images.
Database of Irish Excavation Reports <a href="http://www.excavations.ie/">http://www.excavations.ie/</a>	Summaries of preliminary and final excavation reports
Heritage Data: Vocabularies in a useful form	Monument, temporal period and object vocabularies
ICCU - <a href="http://www.iccu.sbn.it/opencms/opencms/it/">http://www.iccu.sbn.it/opencms/opencms/it/</a>	Bibliographic data
ICCD - <a href="http://www.iccd.beniculturali.it/">http://www.iccd.beniculturali.it/</a>	Archaeological data
Università "L'Orientale" Napoli - <a href="http://www.unior.it/">http://www.unior.it/</a>	Archaeological excavation data
The Cyprus Institute <a href="http://public.cyi.ac.cy/starRepo/">http://public.cyi.ac.cy/starRepo/</a>	Archaeological data
LIBERA; <a href="http://zrcalo1.zrc-sazu.si/libera/lang_en/predstavitev.htm">http://zrcalo1.zrc-sazu.si/libera/lang_en/predstavitev.htm</a>	Libera is a database incorporating primarily archaeological literature. It encompasses the time ranging from the 5/6th century till the 10/11th century in the region of Europe, the Near East and North Africa. The records also include keywords. The database is intended only as a starting point for researching particular problems.)
ZBIVA, <a href="http://zrcalo1.zrc-sazu.si/zbiva/frameset.php?lang=en">http://zrcalo1.zrc-sazu.si/zbiva/frameset.php?lang=en</a>	ZBIVA is an archaeological database for the eastern Alps and its surrounding regions in the early Middle Ages. The test version provides site data with the literature concerning each site
DLIB.SI, <a href="http://www.dlib.si/v2/Default.aspx?&amp;language=eng">http://www.dlib.si/v2/Default.aspx?&amp;language=eng</a>	The Digital Library of Slovenia is linked up with comparable projects in the EU, and is actively involved in building the European Digital Library, Europeana and the TEL project (The European Library).
Monumenta Germaniae Historica, <a href="http://www.dmgh.de/">http://www.dmgh.de/</a>	dMGH are digitised volumes of the MGH edition, made available in Open Access. New volumes are added respecting a copyright term of three years in the dMGH ("moving wall").
Dyabola	bibliographical
Xenon	bibliographical
Persée	bibliographical
Perseus	Ancient texts

DANS-EASY ( <a href="http://easy.dans.knaw.nl">http://easy.dans.knaw.nl</a> )	data archive for Dutch archaeology (e-depot). Scientific datasets published through an online archiving system by various research organisations.
ARCHIS ( <a href="http://archis2.archis.nl">http://archis2.archis.nl</a> )	State service for cultural heritage (RCE): database with research projects, sites and monuments (restricted access)
University of Amsterdam image collection <a href="http://dpc.uba.uva.nl/aac">http://dpc.uba.uva.nl/aac</a> <a href="http://dpc.uba.uva.nl/archeologischecollectie">http://dpc.uba.uva.nl/archeologischecollectie</a>	University of Amsterdam: image collection of original excavation plans of this institute
Specific websites, for example <a href="http://www.narcis.nl/">http://www.narcis.nl/</a>	Wide range of online databases/websites with research data with a limited/specific scope
Library resources, for example <a href="https://openaccess.leidenuniv.nl/">https://openaccess.leidenuniv.nl/</a> or Google scholar	Wide range of library resources and online publications
Archis	Database and GIS-based map system for archaeological research projects, monuments and single finds and observations. Hosted by the Cultural Heritage Agency of the Netherlands (important)
<a href="http://twist.up.npu.cz/">http://twist.up.npu.cz/</a>	Information System on Archaeological Data - database of archaeological excavations and finds for the Czech Republic.
<a href="http://geoportal.gov.cz/web/guest/home/">http://geoportal.gov.cz/web/guest/home/</a>	Map portal providing various map layers.
<a href="http://www.jstor.org/">http://www.jstor.org/</a>	Digital library

We also gave the researchers the opportunity to name further online resources (apart from the archaeological research archives they had listed above) which were important for them in their research, or to make comments. The comments given again confirm the importance but also constraints of the national boundaries (*"... the geographical scope of this database is restricted (Mediterranean), therefore it cannot be used for mainland Europe ..."*), as well as the growing importance of general encyclopaedia and websites also for research purposes (e.g. Wikipedia, Google EarthPro). One of the comments also indicates that not all researchers are necessarily working with raw (primary) data, but may rather rely on published articles.

- *"for data: none; for information: online journals, library catalogues, and even Wikipedia"*
- *"I look at online published articles only (pdf articles)"*
- *"The sources I have mentioned in the above question are not strictly databases, but these are very important online resources for archaeologists. There is also ceraDAT for ceramic analysis that would be particularly useful for me but the geographical scope of this database is restricted (Mediterranean), therefore it cannot be used for mainland Europe: <http://www.ims.demokritos.gr/ceradat/?PHPSESSID=dnjlxknl> "*
- *"Google EarthPro as a virtual landscape measuring distances and directions as well as starting point for archaeoastronomical researches"*
- *"Within the project in which our archaeological data archive at DANS was developed, we have experienced that most of the scholars are especially interested in online access to publications, preferable to a complete set of publications and grey literature reports. This acts as a starting point for further research and provides the opportunity to make a limited selection of resources (sites) to investigate more profoundly and for which we have the time to inspect the original (raw) data."*



### 3.2.4 Shortcomings experienced in using online archives

We then asked the researchers to name the main shortcomings and problems they are experiencing in the use of online archives and databases in their own work. The long list of answers received (see below) is already a clear sign of the many problems that still exist in making research data available online. The problems experienced can, theoretically, be grouped in the following main categories:

- Challenges stemming from a **lack of data quality or metadata quality** (for instance because of missing data);
- Challenges in **having access to data** (e.g. because of restricted access, technical issues how data are stored, or because of cost or copy right issues), and – as a closely related issue – **IPR issues as a barrier** for accessing/using data from online sources
- Challenges due to **problems with data integration** or with **the organisation of the archives** (e.g. difficulty to find and extract coherent datasets).
- **Other challenges**, including specific technical issues.

We found evidence for all these types of challenges in the use of online data sources, as the following examples (extracted statements from the comments of our lead users) demonstrate.

#### Experienced lack of data and metadata quality

- *“No control of data quality – scarce information, no information on data provenance, mostly processed data available, no primary data”*
- *“Low quality of photographs, images, plans etc.”*
- *“Lack of online detailed scientific documentation about objects or monuments that can be used as reference”*
- *“There is a lack of documentation, which makes it time consuming to fully understand the “why” and “how” of specific constructs in the databases.”*
- *“Human made errors are difficult to spot in large data sets.”*
- *“Completeness and precision in geolocation terms”*
- *“Updating of datasets, data quality”*
- *“a lack of metadata quality”*
- *“a lack of consistency in the descriptions”*
- *“relevance of data, completeness of data.”*
- *“Metadata often missing therefore the provenance of the data is not known”*

#### Difficulties in having access to data

- *“Lack of online access to images due to copyright issues”*
- *“Access restricted with commercial conditions”*
- *“Descriptions on how to access the data for machine processing are sparse.”*
- *“In some cases it is difficult to access the actual information, because e.g. it is inside a scanned pdf without OCR.”*
- *“Having full access to datasets”*
- *“Authentication and authorisation procedures”*

- *"The main shortcoming for the moment is still the limited availability. For instance: a specific excavation may look very promising, and I really would like to have the original data to verify their interpretation. But I only have online access to the publication or just a summary of the excavation data. As academic researchers, we would often need to have access to the original raw data and evaluate it (again) ourselves. National sites & monument records only provide a site summary and therefore are only helpful in the initial phase of the research. Access to the underlying raw data has also become vital to the fight against plagiarism and fraud."*

### **IPR issues as a barrier for (re-)using data**

- *"Unclear licensing models or missing licensing lead to problems when it comes to publication."*
- *"Licensing policies, IPR issues"*
- *"copyright uncertainty"*
- *"restrictions at copying for non-commercial personal use"*

### **Problems in the integration of data from different sources / in the organisation of digital archives**

- *"The contributing databases were raised from different research questions. In order to adjust them to the needs of my work I have to pimp the data within a new database (part of the data can be transferred via common interfaces but most details have to be added manually – most times due to the data quality or lack of detail)."*
- *"Data are dispersed in many websites, online databases, portals etc. around the web. It's not always easy to find what you are looking for or to know what is out there."*
- *"Data formats, a lack of standardisation"*
- *"One cannot filter data into periods and geographical area. You receive articles/presentations etc. in your thread on your chosen subjects but you have to decide which is useful for you from the tremendous amount of information (thread) that keep coming daily."*
- *"the standards used"*
- *"Lack of general interface for linked resources."*
- *"lack of standardisation of metadata"*
- *"Data can be quite shallow i.e. point data on a map may tell you what it is but will not take you to a more in-depth account or the relevant data to that object, i.e. lack of linking between datasets"*

### **Technical and other issues**

- *"limited search-possibilities"*
- *"Connectivity, computing and visualization"*
- *"There are no suitable online databases for ceramic analyses for us. The few that are available focus on very specific periods and geographical areas. These would need more "advertising" that encourages researchers to upload their data and widen the geographical area from which data are uploaded. At the moment they are not suitable for cross-regional comparisons except ceraDAT which covers different parts of the Mediterranean."*
- *"Some portals such as Research Gate and Academia.edu do not provide structured scientific data like ceraDAT or other online databases. These provide articles and presentations and not necessarily the whole raw data."*

- *"Data available from recent years only, some restrictions regarding free download, cost issues related to alternative paid-for databases"*
- *"Integrating the data into your own research. Only available if a web feature service or download capabilities"*
- *"Online access to datasets on our national archaeology is limited. Some accessible data is not up-to-date or complete and may contain misrepresentations or inaccuracies."*

### 3.2.5 Sharing/publishing research results online

#### Current practices

We asked our lead users to what extent and in what way the data the research groups were producing in their projects was typically being published, i.e. made available to a certain community beyond their own institutes. We received the following responses (quotes anonymised shown as [...]). They confirm mostly the answers received on a similar question in the Online Survey of 2013.

- *"So far none, the [institute] plans to publish their research databases, they also lead the [xxx] project – an archaeological data service portal, yet being work in progress"*
- *"The data created through projects for ICT applications in culture are published in conference proceedings or in journal articles. Moreover, multimedia archaeological data are made public through project websites or other sites, such as YouTube etc. Raw data, such as excavation reports, finds descriptions, excavation pictures etc. are published in website with limited access."*
- *"In IANUS: via website (cc-by-sa), almost everything we produce. Later on parts of it will be transferred into the online archive (still in work)"*
- *"at the Institute: at national level: Internet Culturale, the portal of the digital resources of the Italian libraries, and CulturalItalia, the national aggregator for Europeana. At European level: ATHENA Plus, EUROPEANA PHOTOGRAPHY, EUROPEANA SOUND, AMBROSIA"*
- *in the [...] Project: constantly, via webGIS specifically dedicated (<http://webgis.archeoroma.beniculturali.it/>); periodically, through monographs, conferences and proceedings, also in open access version."*
- *"2D, 3D and text content published in an open-access repository"*
- *"Data published is primary and processed, with a comprehensive metadata description on the provenance and processing steps of data."*
- *"Articles and reports are published with relevant scientific data and pictures. All raw data available in excel or mdb format just as all the pictures (eg. microphotographs which are not included in the reports). These are deposited in our digital archive and are available for request. These are not available on a website but we would like to provide these through ARIADNE."*
- *"Data is published in form of reports (text) and maps (google earth)"*
- *"In the majority of cases the data is published in catalogues and periodicals."*
- *"Over reliance on the monograph being the primary output from research projects. Now within new projects formalised data management strategies are being developed in conjunction with the cultural heritage sector in Ireland to formalise data deposition, archiving and access following a research project "*

- *“Book and journal publishing; online through dedicated websites, portals such as Europeana, and through infrastructures provided through projects such as 3D ICONS and LoCloud.”*
- *“The results of a research project are made available to a wide community by data archived into our public repository, data published on online journals and data referred by social networking website for academics.”*
- *“paper documentation digitalized in PDF, data-bases in table view in some very common format (Excel, Access ...), printed monograph (available as PDF, too), all of above free available over internet, Interim reports on Fasti Online Documents and Research, or on Project's website, Invited lectures, talks, etc.”*
- *“Full publication in print, with complementary website of catalogues and primary data.”*
- *“Within the academic environment publication of research data is limited. Reports are made available in digital form through local repositories. In the environment of (commercial) contract archaeology it is required to deposit all data and documentation that is produced during research in [country] at an (electronic) depot within two years after completion (Malta legislation).”*
- *“Scientific publications, website of a project”*

### Archives and repositories used for depositing data

We also asked the researchers if they were (at least occasionally) depositing research data from their research projects in an online database (other than their internal institutional database), which can be accessed by other researchers. We asked those that did to name the most important archives/repositories where they had deposited data in the past. The following sites were mentioned:

Archive	Focus/description	Website
EBIDAT	Medieval castles	<a href="http://Ebidat.de">http://Ebidat.de</a>
Europeana	Europe's culture in digital format	<a href="http://www.europeana.eu/">http://www.europeana.eu/</a>
STARC - Repo	Internal Data	<a href="http://public.cyi.ac.cy/starcRepo/explore/objects">http://public.cyi.ac.cy/starcRepo/explore/objects</a>
Academia.edu	network of researchers, depository of published and unpublished works	<a href="http://www.academia.edu">www.academia.edu</a>
ResearchGate	network of researchers, depository of published and unpublished works	<a href="http://www.researchgate.net">www.researchgate.net</a>
cimec	National archive	<a href="http://www.cimec.ro/">http://www.cimec.ro/</a>
Mapping Death	Burial and cremation data	<a href="http://www.mappingdeathdb.ie/">http://www.mappingdeathdb.ie/</a>
Wodan	Paleo-wood & Charcoal data	<a href="http://wodancharcoal.ie/">http://wodancharcoal.ie/</a>
www.3dicons.ie	3D survey data for cultural heritage	<a href="http://www.3dicons.ie">www.3dicons.ie</a>
ADS Archaeology Data Service	Preserving and disseminating digital data in archaeology	<a href="http://archaeologydataservice.ac.uk">http://archaeologydataservice.ac.uk</a>
3D – Icons	Archiving of digital archaeological resources	<a href="http://vast-lab.org/3dicons/index2.php">http://vast-lab.org/3dicons/index2.php</a>
Europeana	Archiving of digital resources	<a href="http://www.europeana.eu">www.europeana.eu</a>
DLIB.SI	described above	<a href="http://www.dlib.si/v2/Default.aspx?&amp;language=eng">http://www.dlib.si/v2/Default.aspx?&amp;language=eng</a>
Sistory	Slovenian history	<a href="http://www.sistory.si/?language=en">http://www.sistory.si/?language=en</a>
Fasti Online/FOLD&R	Excavation database	<a href="http://www.fastionline.org">www.fastionline.org</a>

Excavation Dronten	dataset	<a href="http://www.persistent-identifier.nl/?identifier=urn:nbn:nl:ui:13-ntxf-4n">http://www.persistent-identifier.nl/?identifier=urn:nbn:nl:ui:13-ntxf-4n</a>
Excavation Schipluiden	dataset	<a href="http://www.persistent-identifier.nl/?identifier=urn:nbn:nl:ui:13-dg3-qfx">http://www.persistent-identifier.nl/?identifier=urn:nbn:nl:ui:13-dg3-qfx</a>
Excavation A27 - Hoge Vaart	dataset	<a href="http://www.persistent-identifier.nl/?identifier=urn:nbn:nl:ui:13-0pr-of8">http://www.persistent-identifier.nl/?identifier=urn:nbn:nl:ui:13-0pr-of8</a>
Europeana:	Resource for European Cultural Heritage. Open Access archaeological publications from EASY are harvested and made accessible in Europeana (enabled within the CARARE project).	<a href="http://www.europeana.eu/">http://www.europeana.eu/</a>
Archeological database of Bohemia	the central database of archaeological excavations and finds for Bohemia	

### 3.2.6 Other issues

Finally, we gave the lead users an opportunity to provide any further information or feedback with regard to their research activities which was not covered by any of the questions about the specific portal they had evaluated. We received the following comments:

- *“If datasets are available through portals as web services, it is crucial to know that these are persistently moving forwards.”*
- *“There is a lack of infrastructural support for depositing archaeological research data in [country] which provides for timely accessibility and usability of that data. Where data is accessible online, a website is often built with the research project or dataset to provide access.”*
- *“Many researchers are not yet skilled in using and analysing existing online databases.”*
- *“The archaeological data archive in [country], as currently available at [portal], is still limited in its research functionality. This archive is a good starting point, simply safeguarding most of the digital data produced during an archaeological research project. The way digital information can be searched, founded and integrated could be improved as is shown by many modern online systems. A full text search through all archaeological publications or cross-searching the data of multiple project archives, are just two examples that would help archaeologists in their work.”*

## 4 Evaluation of archives and portals

### 4.1 The sample: archives and portals evaluated

The sample studied to identify best practices consists of 25 research archives and portals. Some archives/portals were looked at independently by several researchers in parallel (Arachne, Global Change Master Directory, Pleiades), so the number of questionnaires received is higher than the number of portals/archives (the responses were collected through 26 questionnaires returned to the study team; the IDs in the following overview refer to the questionnaires).<sup>5</sup> In this section, we introduce the 25 archives and portals in more detail.

In terms of their geographic focus and scope, two thirds (16) of the archives and portals held global collections or did not have a specific regional focus; four archives were European-focused and five had a national focus (see Table 4.1).

From a thematic/domain perspective, nine archives/portals were specifically established to provide information about general archaeological research (see Section 4.2); five have a focus on specialist archaeological domains (e.g. with collections on cultural heritage, history, arts; see Section 4.3), eight archives/portals are about other scientific domains (mostly natural sciences, see Section 4.4) and three are for research communication in general without any domain focus (see Section 4.5). This is an indicative classification, though; the borders, in particular between general and specialist archaeological research, are not always easy to draw.

It is even more challenging to group our sample into different types of archives or portals, even if their characteristics differ significantly in terms of the services that are offered. Some initiatives focus predominantly on providing access to data collections, with few additional services, such as guidelines or collaboration tools. Others could be described as extended “research portals” – their services go beyond the holding of data sets and collections; they provide additional information and services, for instance guidelines, link to other organisations and institutions, or offer exchange mechanisms. Finally, we have included three portals in our analysis where collaborative support within the research community through specific services is in the foreground (rather than the collection and provision of data in the narrow sense).

**Table 4-1: Structure of the sample of portals/archives**

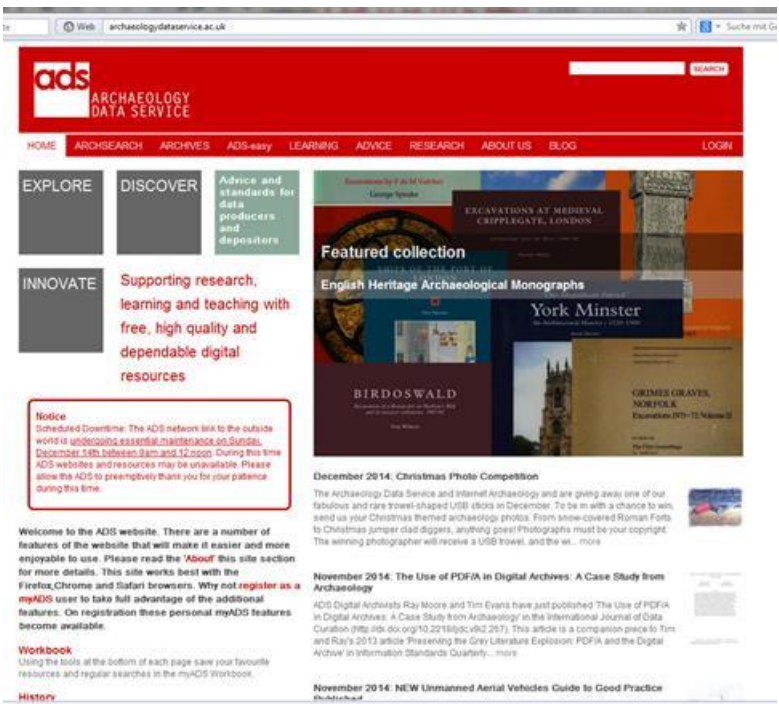
Regional focus	No.	Thematic focus	No.	Type	No.
National	5	General archaeology	9	Database/archive	15
European	4	Specialist archaeology	5	Research portal	7
International/global	16	Other domains	8	Collaborative portal	3
		Research communication	3		
<b>TOTAL</b>	<b>25</b>		<b>25</b>		<b>25</b>

<sup>5</sup> We decided not to publish in this document the names of the individual researchers who carried out the assessment, but to provide only the names of the organisations from which they were selected for this purpose. Annex II offers an overview of the archives and portals and the organisations from which the evaluators were selected.



## 4.2 General archaeology portals

### 4.2.1 ADS – Archaeology Data Service

Website	<p><a href="http://archaeologydataservice.ac.uk">http://archaeologydataservice.ac.uk</a></p> 
Provider	The ADS was founded by a consortium comprising the Council for British Archaeology with the Universities of Birmingham, Bradford, Glasgow, Kent at Canterbury, Leicester, Newcastle, Oxford and York. The ADS is guided by an advisory committee consisting of representatives from all sectors of the discipline. The ADS is based at the University of York.
Domain/thematic focus	British archaeology, excavation reports
Geographic focus	UK
Type of data offered	Excavation data, reports, publications, GIS, statistics, natural science data
Further information	The ADS aims to support research, learning and teaching with freely available, high quality and dependable digital resources. It does this by preserving digital data in the long term, and by promoting and disseminating a broad range of data in archaeology. The ADS promotes good practice in the use of digital data in archaeology, it provides technical advice to the research community, and supports the deployment of digital technologies.
Questionnaire ID	01

## **Assessment of the portal – strengths, weaknesses, innovative features**

The evaluator specified a hypothetical research task for carrying out the assessment of the portal's functionalities consisting in doing research for a project on Iron Age hillforts in the sphere of a Roman Frontier (a comparative study using examples from UK and Germany). The research questions were (e.g.) what happened to indigenous people in these areas, and how did land use and habitation change with the arrival of Romans.

### **Overall structure of the portal**

The overall structure of ADS was found to be good. Everything is well structured, services can easily be found. The web page looks "tidy and modern", as the user noted, and could thus serve as a good practice for ARIADNE as a whole. The user did not encounter any specific best practices, however.

### **Search functions**

The user experienced the search mechanisms as rather sophisticated in the beginning. There are different entry masks and different levels of specification. From this point of view, he rated the search as "good". The problem was, however, that in the overall search of Iron Age and Roman domestic sites in the middle of the UK, the user expected a wide range of hits within different districts. When he tried to tip different counties and/or districts within the UK, the search results were limited to an intersection of reports containing all the items he had selected, but he wanted the total volume of all reports within the selected areas depicting iron age, roman and settlement. This was embarrassing, for one had to depict each geographic region individually.

The overall selection within the advanced search offer was good, however.

### **Mechanisms for downloading data**

The download mechanisms were found to be good. One can download search results in xml and csv format. Reports and other data can be downloaded from the individual record for each report in the webpage given its use is not restricted.

### **Options for uploading data**

The procedures for uploading and sharing data seem to be well organised, but could not be practically tested in this evaluation. The portal tells you for which data and formats the archive is suitable for.

### **Specific support functions & features offered**

The portal offers a rich selection of information materials, guides, teaching and training materials, feeds and newsletters.

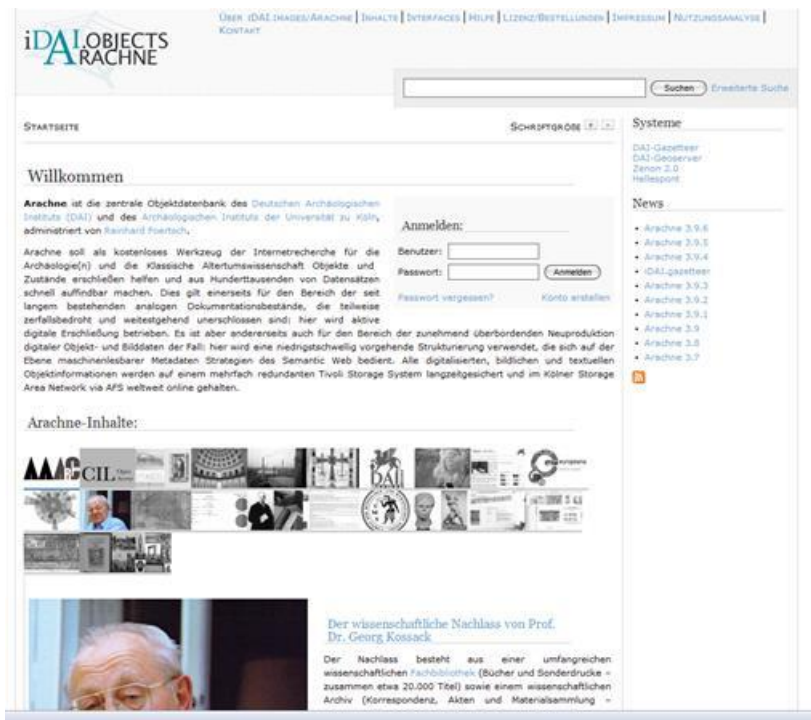
## **Conclusions for ARIADNE**

The following suggestions from the evaluation of ADS by a lead user could be taken into consideration in the design of the ARIADNE portal:

- Good practice: the overall structure and organisation of the portal
- Good practice: the rich content of the portal (on top of the data available), in particular the availability of guidelines.



## 4.2.2 Arachne

Website	<p><a href="http://arachne.uni-koeln.de">http://arachne.uni-koeln.de</a></p> 
Provider	University of Cologne, Archaeological Institute German Archaeological Institute
Domain/thematic focus	Arachne is the central object database of the German Archaeological Institute (DAI) and the Archaeological Institute of the University of Cologne, providing a free internet research tool for quickly searching hundreds of thousands of records on objects and their attributes. It is one of the major archives/portals for archaeological research in Europe.
Geographic focus	Europe, Asia and South America. Arachne initially started with a focus on the Mediterranean, but has expanded significantly.
Type of data offered	The focus is on object data, such as sculptures, building parts, coins, and seals. There are also plenty of digitised photographs and texts.
Further information	Arachne has been operating since 2004. The initiative aims to foster interoperability between different systems while at the same time protecting copy rights of the content owners.
Questionnaire IDs	02, 03, 04

## Assessment of the portal – strengths, weaknesses, innovative features

The three evaluators specified hypothetical research tasks for carrying out the assessment of the portal's functionalities.

- Lead User 1 aimed to search and find the data available from ancient Thessaloniki. The lead user wanted to see if it would be able to download the data and in what format.
- Lead User 2 reported that his hypothetical task was based on his actual experiences while doing his research, including: searching, downloading and preparing the data for further tasks.
- Lead User 3 conducted research on Augustus Imperator for a TV documentary and was looking for bibliographic and iconographic sources.

### Overall structure of the portal

The lead users (LU) testing Arachne rated the overall structure of the portal as "good" (LU 1, 3) and "average" (LU 2).

Lead User 1 notes: "It is a site containing a huge number of digital data that is rather well organised. The objects are divided in collections but one can browse all collections using the search box. It has few and easily understood pages/sections. It is easy to navigate since it provides the user path on top of each page. The amount of information displayed on each page for an object or monument is not the same for all items, it can be very few or more." The user highlights a good practice in the structure: the "Help" section is displayed always on the right side of the page and it is guiding the user at all pages. This was found to be a very helpful feature.

Lead User 2 notes that the site looks "cluttered" due to the small default font size and the placement of the different navigation elements. They (and LU 3) criticise that the language was set automatically; there was apparently no opportunity to change the language (e.g. toggle between English and German). They appreciated the display for search results, however. Further information is displayed on the right hand side when hovering over single results or some navigation elements. This was found to be useful. Further good practices of the portal's structure include:

- dedicated browsers for single projects (e.g. Corpus of Minoan and Mycenaean Seals),
- providing index and value lists for searches, and
- providing dedicated browsers for specific projects

They suggest that the overall documentation could be expanded (especially for "interfaces" and the "help" section).

Lead User 3 provides extensive feed-back on the structure. They find that the design of ARACHNE is generally good, as the web pages are clear and the contents are well presented. They note, however, that the accessibility has to be enhanced to meet the requirements of WCAG 2.0 and gives specific examples.

Regarding the geolocation of the resources, the DAI.gazetteer offers a first attempt for georeferencing the objects, but it could be improved with a more detailed cartographical representation of all those objects for which it is possible, e.g. using the "DAI-Geoserver" accessible from the ARACHNE portal.

The structure of the metadata resources are described in three different models: METS, MODS and TEI. The resources are made available via OAI-PMH (<http://arachne.uni-koeln.de/drupal/?q=en/node/235>). The ARACHNE dataschema is mapped into CIDOC-CRM to make the semantic contents "machine readable". Thanks to the mapping into CIDOC, each digital object is identified geographically (by place of origin ex: Kerkyra, Corfù), importing the ontology defined the gazetteer iDAI.

However, despite the fact that resources are accessible in a satisfactory way, there is a lack of standardisation in categorising the resources, using a metadata structure developed "ad hoc" rather than an approved ISO series.

For the visualisation and the indexing functions, the system is based on SOLR which indexes the CIDOC-CRM ontology and displays the facets. Technical documentation, available on the web site, should better describe how it works within the ARACHNE system. In fact there isn't a section devoted to technical documentation about the portal and the interoperability, or any guidelines on digitization for the archaeologists.

The visualization tool has been developed specifically for the project: while looking into the archaeological information provided by ARACHNE, a user might select one particular archaeological object and then, they can switch to a tool that is called the "Context Browser". It visualizes links to additional information being related to the object that is currently on the screen and it provides information about the context of a find and defines additional contexts like affiliation to a specific collection or to a historical monument. The viewer was found to be not state of the art, and not very usable; there are other frameworks (d3js, nodejs, bokeh, etc.) that allow the semantic navigation in a more intuitive way.

LU 3 suggests that, regarding the data licenses policy, users should be informed about re-use conditions applied to each resource before they access a purchase or contact procedure. Currently, it is reported: "...Contents that can merely be accessed in an indirect connection with our web site Arachne (e.g. high-resolution printable scans) are subject to different regulations that do not pertain to the Creative Commons License. These regulations are explained in context with the transaction of images after a purchase or in the contact form."

The user also highlights that the user registration procedure is not very efficient: "After having filled in the form, I received the confirmation e-mail which was sent automatically by the system, but the authentication procedure was only completed after a second email, sent two days later by project staff."

The Lead User reported the following good practices in the structure:

- The semantic organisation of data based on the CIDOC-CRM ontology: "It is one of the few examples of implementation of CIDOC-CRM in the field and, therefore, constitutes a precedent.
- The possibility to view and browse composite images, like in the "Monuments browser" section: it is possible to browse surfaces in a continuous plain view (e.g. the Trajan's Column) and search iconographic details and entire scenes.
- The metadata download procedure: the download is very easy and intuitive and the resources are already formatted in XML. However, within the XML only the TAGs are shown related to the resource examined, without any RDF relations that link the resource to the others and that could be a limitation.

### Search functions

The search functions which the portal offers were rated as "good" (LU 1, 3) and "average" (LU 2). Lead user 1 reports that it is "rather easy to find data since you only have to fill in a word in the search box available in the home page". The portal has an extended search function available with many optional criteria that can prove to be useful, too. It was found useful that one can carry out a thematic search first and then continue by selecting on the basis of different criteria. A limitation is that the extended version is only available in German. This is also the case for other sections of the website.

Lead user 2 describes their experience with the search function: "There is a basic search and an extended search. When using basic search, an overview with number of hits for each category is shown, which then has to be selected. For the extended search, an individual context has to be chosen (the image provided for choosing and the additional help texts are good)."

They miss the option to search with "or" and similar search operators.

They also suggest that an actual "search" button to start the search would be useful, rather than having to press enter.

The provided index and value lists were seen as helpful, but an option was missed to narrow down current search results. The user also suggested that highlighting the matching search keyword would be nice.

A good practice was seen in the option to display search results in three different formats (images, lists, catalogues) and the option to choose the number of displayed results per page.

Lead User 3 notes that search functions are rather efficient. It was noticed that full text search retrieves rather different results for synonyms of a specific term (for instance, "colosseum" for "colosseo" or "giove" for "iuppiter", "atene" for "athene", and viceversa). They suggest the search could be enhanced through a categorization based on normalised multi-language thesauri, and the possibility of contributions from the user community could be considered.

### **Mechanisms for downloading data**

The mechanisms for downloading data are seen as "average"/"fair". One of the users noted that it was easy to download data, as there was a "disc" icon on every item page (objects, monuments etc.) that directed the user to the available download formats. One can download data in different formats, including Microsoft Word, pdf, xml or html. This was found to be useful.

The other user noted that there was no direct download button, nor a basket to store items and download them as bulk. The user also stated: "Downloading data on Arachne can be done via provided APIs, e.g. via OAI-PMH. But this is something not everybody is willing or able to do." He also missed documentation where e.g. information is provided about how to address the server.

### **Specific support functions & features offered**

The user noted the following support functions and special features:

- Community services: the possibility to create your own portfolio of images and to log in to have access to more high quality images.
- Flexibility of tools (e.g. mobile versions, apps): the possibility to increase/decrease font sizes.
- Alerting/news feeds: a news section, and feeds only available in the home page.
- Thematic collections: specific projects have their own interface, which allows for browsing in the project material.
- Linking to other services: places and literature data contain links to GeoNames, DAI-Gazetteer and Zenon.

One of the users notes that the site does not take fully advantage of social networking, which they would have appreciated.

A good practice was seen in the availability of open data.

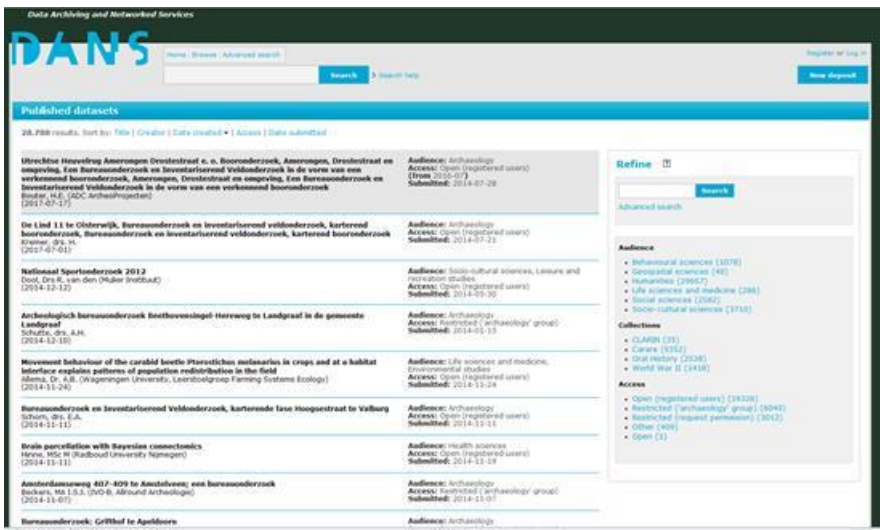
A specific suggestion was to adjust the design of the page so that it also fits on tablet devices, while a dedicated app was not seen as necessarily needed.

## Conclusions for ARIADNE

The following suggestions from the evaluation of Arachne by a lead user could be taken into consideration in the design of the ARIADNE portal:

- Good practice: displaying further information on the right hand side of a page when hovering over single results or some navigation elements.
- Good practice: The semantic organisation of data based on the CIDOC-CRM ontology.
- Good practice: the metadata download procedure (download is very easy and intuitive and resources are already formatted in XML).
- Good practice/useful feature: a “Help” section which is displayed always on the right side of the portal and guides the user through all pages.
- Good practice: display of search results in different formats (images, lists, catalogues).
- Helpful feature: if different languages are offered, there should be a button to switch between languages on a page by page basis (e.g. toggle between English and German).

### 4.2.3 DANS EASY – The online system of the Data Archiving and Networked Services

Website	<p><a href="https://easy.dans.knaw.nl/ui/home">https://easy.dans.knaw.nl/ui/home</a></p> 
Provider	DANS, an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organisation for Scientific Research (NWO)
Domain/thematic focus	Cross-disciplinary humanities and social sciences research data, in particular from the following domains: history and oral history, archaeology, social and behavioural sciences, geospatial sciences. In addition, there are “thematic collections” of datasets that are based on the same research theme (e.g. Data Education Research, Data Youth Research).
Geographic focus	Global (the thematic collections are mostly available in Dutch only)
Type of data offered	Scientific articles on searched subjects – results of data analyses (access limited to registered users)
Further information	<p>EASY is the online archiving system of Data Archiving and Networked Services (DANS). EASY offers you access to thousands of datasets in the humanities, the social sciences and other disciplines. EASY can also be used for the online depositing of research data. A majority of the datasets are Open Access available.</p> <p>A large number of protected micro-data held by Statistics Netherlands (CBS) is also available free of charge, via DANS, to researchers at institutions granted the relevant authorisation under the Statistics Netherlands Act. CBS data can be delivered to students who are officially employed by their university.</p>
Questionnaire ID	27

## Assessment of the portal – strengths, weaknesses, innovative features

The hypothetical research task chosen by the evaluation team was related to native settlement patterns from the Iron Age to the early Roman period. The goal was to look for contributions covering the Iron Age, Roman settlements and landscape studies. These comprise works with biometrical and geo-archaeological content. Ideally, the research team hoped to find coordinate data (csv) that enable mapping and some statistical analyses.

### Overall structure of the portal

The evaluators appreciated the clear and simple structure of the portal. For searching and browsing the available information, one can choose between different filters or searching for certain topics (using a simple or advanced search mechanism). It is also possible to deposit data and to create and administer a user account.

### Search functions and mechanisms for downloading data

The search function was rated as “average to good”. The evaluation team notes that the advanced search does not offer many options. If one does not look for specific datasets, the only options for filtering/classifying information are the “any field” or “title” search.

For this type of search, the browse options are much more relevant and interesting. Acknowledging this drawback, it is at least intuitively clear which options are available and how to use them. There is also a useful help function, explaining the different options for a combined search. The function for browsing has a logical structure and offers sophisticated refinements. For archaeological datasets, there is a thorough chronological classification, and one can look for types of sites or sociological items.

The search process itself was found to be quick; it does not take much time for processing and delivering the results. There is also a possibility to refine/narrow down the search (e.g. browse in specific categories) once the initial search results have been delivered. This was found to be useful.

However, one of the evaluators notes: *“It would have been of use for my particular search topic to be able to select multiple items/terms within the “browse” function.”* The evaluators also missed an option to personalise search processes and results, for example by being able to save specific search results or combinations of search filters (“my search”).

Downloads are available from the description table in EASY. One can view data files provided by the author and download them, provided that permission is given. The download process is easy to accomplish. Meta data are provided as xml or csv files. Additional data is mostly provided in pdf format. In some cases, data include tables and/or images.

### Options for uploading data

The mechanism for depositing data was found to be easy and well structured. Researches who wish to deposit data need to be authorised. To obtain this, a researcher has to belong to a registered research group or institute. The evaluators appreciated this procedure, as it was seen as a quality assurance measure. The uploading process for depositing data was found to be simple, and the handling of IPR issues was experienced to be transparent and well structured. All in all, the simplicity and straightforward structure of the data deposit procedure was recommended as a good practice and possible role model for other portals.

### Specific support functions & features offered

The DANS-EASY database itself does not provide additional information or services, but users can find further information on the main DANS website. An evaluator remarks: *“I think it is not necessary to have thematic collections etc. within a research data portal, because archaeologists tend to come*



*with specific research issues they expect to retrieve. Maybe it would be nice to have find spots displayed on a map, but this is not an academic necessity."*

It was suggested (not for DANS-EASY in particular, but for research portals in general) that a review system could be helpful to further assure data quality within portals; however, there was a concern that such a procedure, while improving the quality of the datasets, might at the same time be a barrier which would make depositors less willing to share data, as it makes the process more complicated.

### **Conclusions for ARIADNE**

The following suggestions could be taken into consideration in the design of the ARIADNE portal:

- Good practice: An option to search in several steps, i.e. have the possibility to refine or narrow down the search parameters after an initial, broader search has been carried out, on the basis of the (initial) search results delivered.
- Requirement: Portals should give the option to save specific search results and combinations of search filters ("my searches").
- Good practice: the simplicity and straightforward structure of the data depositing procedure.

## 4.2.4 Fasti Online

Website	<p><a href="http://www.fastionline.org">http://www.fastionline.org</a></p> 
Provider	AIAC – the International Association of Classical Archaeology, and CSAI – the Center for the Study of Ancient Italy of the University of Texas at Austin.
Domain/thematic focus	Classical archaeology
Geographic focus	Balkan Peninsula, Italy, Spain, Ukraine, Morocco
Type of data offered	Excavation reports – mainly of excavations since 2000
Further information	<p>Fasti Online is a project of the International Association of Classical Archaeology (AIAC) and the Center for the Study of Ancient Italy of the University of Texas at Austin (CSAI). The project is supported by the Baron Lorne Thyssen, the Italian Ministry for Cultural Heritage and Activities, the Society for the Promotion of Roman Studies, the Classical Association and the Oxford Journal of Archaeology. Fasti Online emerged from the “Fasti Archaeologici”, published between 1946 and 1987 by the International Association for Classical Archaeology (AIAC).</p> <p>The aim of the site is to provide a database of excavations since 2000, providing a record in English and in the local language for each season. All Fasti Online data, unless otherwise stated, is licensed under the Creative Commons Attribution-ShareAlike Unported License (CC BY-SA 3.0). Searches can be made by country, site name, site date range, monument types, site directors and site locations.</p>
Questionnaire ID	11

## **Assessment of the portal – strengths, weaknesses, innovative features**

The evaluator said they had already used Fasti Online for previous research. They re-used this experience as a “research assignment” for carrying out the evaluation. The assignment included searching the database using keywords regarding chronology in order to check all uploaded excavation reports, and gather the data needed for the respective research.

### **Overall structure of the portal**

The user rated the overall structure of the portal as good. They consider the portal to be well designed and very user friendly. “There is a variety of search filters which makes the navigation easy. The information provided is adequate, and although there are some gaps regarding bibliography and images of some sites, the database is a good starting point for further research. The data is available in local languages but also in English which makes it accessible to wider audiences.”

They regard the clear display of chronology and monument type to be a good practice of this portal.

### **Search functions and mechanisms for downloading data**

The search functions are rated as very good, since one can use different search criteria and, therefore, quickly find the available data. The variety of different search filters (by data range, monument types, country, site name, site directors and site location) is seen as a strength and good practice of this portal.

The downside of Fasti Online is that the data is available only for printing. “One can download only the form for uploading a record.”

### **Options for uploading data**

The upload mechanism was found to be easy. It requires sending a record form to the relevant country administrator who is responsible for checking the quality and legitimacy of the data. This probably implies restrictions to the upload mechanisms, i.e. only specialists in the area may upload data.

### **Specific support functions & features offered**


Fasti Online is focused on the core function and does not offer many added-value services on top. A community service is to inform about opportunities for volunteers to participate in a research project.

## **Conclusions for ARIADNE**

The following suggestions could be taken into consideration in the design of the ARIADNE portal:

- Good practice: the variety of different search filters (by date range, monument type, country, site name, site directors and site location)

## 4.2.5 MAPPA

Website	<p><a href="http://mappaproject.arch.unipi.it/?lang=en">http://mappaproject.arch.unipi.it/?lang=en</a></p> 
Provider	Pisa University
Domain/thematic focus	Archaeological data
Geographic focus	Italy
Type of data offered	Maps, historic maps, aerial and satellite photographs, geomorphological data, archaeological restrictions, cadastral maps, DTM, orthophotos, transportation network, paleogeographic data.
Further information	The MAPPA project aims to create the first Italian open digital archaeological archive that makes public data relating to archaeological investigations accessible. MAPPA shall become a network of systems and standardised procedures for drawing up and handling archaeological data. Archaeologists, geologists and mathematicians shall combine, through MAPPA, their expertise to study predictive calculation instruments applied to the archaeological potential of an urban area.
Questionnaire ID	13

## **Assessment of the portal – strengths, weaknesses, innovative features**

The evaluators say they have explored the general capabilities of the portal to identify the main features provided. In order to better understand the functionalities of the portal, they carried out hypothetical research with a query to find information concerning the roman ships found in a specific area known to them. Their main conclusion was that the discovery of data was too complex and being unable to perform free text searches made data discovery a very long process.

### **Overall structure of the portal**

The evaluation team rated the overall structure of the portal as fair. The content is available only in one language (Italian) except for the homepage. The GIS technology used is quite old and the browsing and search is quite slow, in comparison with other similar services. Notwithstanding the filter criteria is well organized, but the results are not clearly presented and were difficult to understand.

The filter criteria of query results could be considered good practice for other portals.

### **Search functions**

Data is not well organized for the reasons described above.

### **Mechanisms for downloading data**

There is only an option to download images and a csv file with poor information.

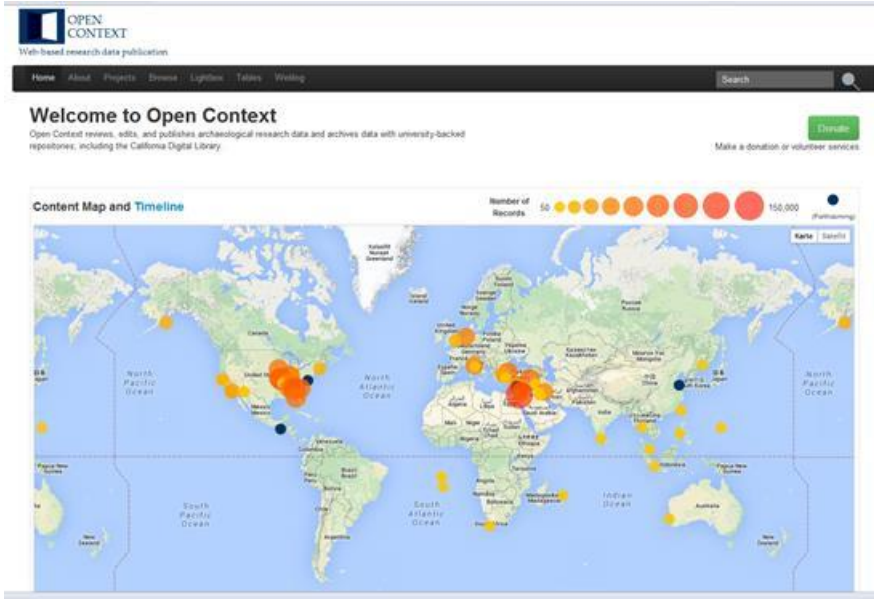
### **Options for uploading data**

The upload/depositing mechanisms are not web-based. There is no web interface to compile. Instead, it is necessary download two pdf files and to send them through e-mail.

## **Conclusions for ARIADNE**

- Good practice: The filter criteria of query results

## 4.2.6 OpenContext

Website	<p><a href="http://opencontext.org">http://opencontext.org</a></p> 
Provider	The Alexandria Archive Institute, a non-profit organisation.
Domain/thematic focus	Archaeology (excavation projects), zooarchaeology
Geographic focus	Global
Type of data offered	Excavation and survey data, some scientific data (micromorphology, archaeometry), zooarchaeological data, Linked Data
Further information	<p>Open Context reviews, edits, and publishes archaeological research data (contributed by researchers) and archives data with university-backed repositories, including the California Digital Library. Thus, Open Context provides a platform for researchers to publish their primary field data and documentation. Open Context licenses all content with Creative Commons, and makes it available in a variety of machine-readable formats. It is a free and open access service, all members of the public are welcome to use and reuse the content. The database claims that its data publications can complement and enhance conventional publications through comprehensive dissemination and preservation of rich digital data and media.</p> <p>The database has currently more than 910,000 items. Searches are possible by region, project (in different domains) or category (e.g. "site", "animal bone").</p> <p>Open Context participates in the " Linked Open Data" community by publishing data linked to URI-identified concepts published by others.</p>
Questionnaire ID	16

## Assessment of the portal – strengths, weaknesses, innovative features

The evaluation team specified a search for Neolithic (and Early Bronze Age) data from Anatolia and Greece, especially excavation data, and tried to find data in order to compare and link them with their own data.

### Overall structure of the portal

The user rated the overall structure of the portal as “good”. The meta-structure was found to be clear; one gets an overview of what is available, can browse and search data, and the site offers a map and a timeline with all data. However, there are some fields where users may not immediately know what they mean, such as “lightbox” (which seems to be a collection of images) or “tables” (downloadable data tables from different projects). The detailed structure when browsing in projects is different from project to project, and there is different information available for various projects. Some are more detailed than others. Overall, information about a project in general is via the project tab, which is not obvious from the beginning.

A good practice in the structure is that most of the data are linked with each other. This is helpful for browsing from a general topic (site) to a specific one (drilling down to loci/contexts and finds). It makes it easy to find out which context contains which finds.

### Search functions

The search functions of the portal are rather limited. There is only one window for searching, no advanced search is possible. Searching with many keywords does not seem to be reliable. Similar queries lead to different results.

### Mechanisms for downloading data

One can browse inside the projects and then find a link to download data, but this is complicated and hard to find. The site offers a tab called “tables”, which lists the projects which have data for downloading (as csv). Unfortunately, not all links are working, and there is no information available for the fields of the table (no explanation what they mean, what the content of the fields is).

The downloading process itself is easy and worked well.

### Options for uploading data

Users cannot upload data by themselves. They have to write an email or fill in a form to deliver some information about their data first. After this procedure, they will be enabled to upload (and link) their data. It was suggested that open formats should be used for this. Fees vary between USD 250 - 6000, depending on the size and complexity of the dataset to be shared.

### Specific support functions & features offered

There is an extra website with a Weblog that provides information e.g. about conferences, data publications, editorial workflow, events, news and policy developments. This Weblog is quite elaborate, but seems to not be used very often. One can leave comments, but there are very few comments posted. There are no specific tools (e.g. semantic web tools or similar), and the last entries are dated from September 2014.

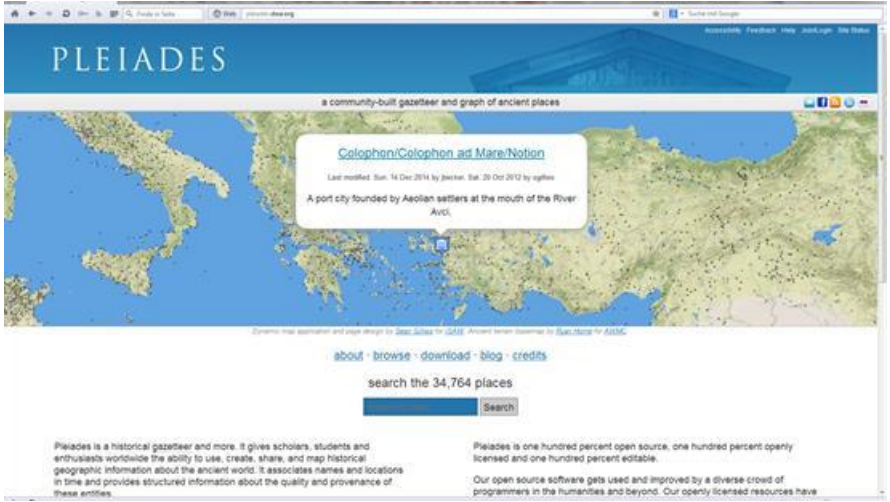
## Conclusions for ARIADNE

The following suggestions could be taken into consideration in the design of the ARIADNE portal:

- Good practice: The use of Linked Data. Most of the data is linked with each other. This is helpful for browsing from a general topic (site) to a specific one (going “down” to loci/contexts and finds) and makes it easy to find out which context contains which finds.



### 4.2.7 Pleiades

Website	<a href="http://pleiades.stoa.org">http://pleiades.stoa.org</a> 
Provider	The National Endowment for the Humanities, an independent US federal agency created in 1965.
Domain/thematic focus	Ancient geography
Geographic focus	Mediterranean and Europe as part of the classical world
Type of data offered	Localization of ancient point data and provision of basic references. The link to Pelagios adds some, though not all, of the ancient sources for the place. In turn, this links the place to Arachne, which holds various published images for the site.
Further information	<p>Pleiades calls itself a “historical gazetteer and graph of ancient places”. It aims to give scholars, students and interested parties worldwide the ability to use, create, share, and map historical geographic information about the ancient world. It associates names and locations in time and provides structured information about the quality and provenance of these entities. The platform has entries to about 35,000 ancient world “places” (geographical and historical contexts for names and locations).</p> <p>Names and locations are collected into conceptual bundles (places) and these collections are associated with other geographically connected places. The catalogue of Pleiades objects can be searched via simple or advanced forms and is written every morning to files that can be opened in a spreadsheet programme.</p> <p>All holdings of Pleiades are open source, openly licensed and editable.</p>
Questionnaire ID	18, 19

## Assessment of the portal – strengths, weaknesses, innovative features

Pleiades was evaluated by two lead users.

The hypothetical research assignment for LU 1 to carry out the assessment was to search and find the geographic data available for ancient Abdera, and to check if there is reference to the change of names for this ancient Greek city and other locations with the same name. Then, the evaluator wanted to see if it was possible to download the data, and in what format.

LU 2 was highly familiar with the portal and shared some general considerations. LU 2 remarked that the site was “oddly ill-suited to actual research” for specific reasons. The portal is not so much a database in the strict sense, but rather a framework on which to hang other information, “effectively a gazetteer”. It is thus best for looking up individual places, and indeed its stated aim is to establish stable, unique identifiers for ancient places and information about them. Having concluded that it was not usable as a standard “sites and monuments’ database”, LU 2 decided simply to examine it for sites they knew well, to see what information was there, and how accessible it was. The creators of the site argue that the point is to use it as a building block for a disambiguated map of ancient places, onto which structure linked data can be connected. This is also the aim of the sister site, Pelagios, which connects occurrences of toponyms in texts to those of Pleiades, with further connections to Arachne, CLAROS, DARMC, Perseus (but not Persée) and other resources.

### Overall structure of the portal

LU 1 rated the overall structure of the portal as “average”. It is a simply designed site, with few and easily understood pages/sections. It is easy to navigate, as it uses maps and links to the places of interest. The amount of information displayed on each page for a geographic position is rather modest. Apart from the geographic data, at least a short summary describing each place would be useful.

It was difficult for the user as an archaeologist to understand that “KML” stands for data viewable on Google Earth (see above).

A good practice is the provision of content in the form of maps and coordinates, and also the references and links to related content in other online datasets (Pelagios).

LU 2 describes the structure of the site. The home page gives five tabs Home, Browse Places, Download Data, Blog, Credits, Documentation, Vocab. There is a map of recently modified resources, and a quick ‘about’, as well as a column for ‘news’. The rest of the tabs are self-explanatory. The evaluation of the structure is as follows: “Within its terms it is good, though the lack of a map that can be interrogated (except by downloading the KML file to Google Earth, and then attempting to click on the resulting points, which will then return you to Pleiades) makes it a gazetteer and not an atlas.<sup>6</sup> Interrogating a site name brings up a choice of options, clicking on the appropriate one brings up the position of the site, and, on a table to the right, links to the various databases that carry information about it. Pleiades itself carries only the information from the Barrington Atlas, and not always that (as in the case of the ‘agglomerations’). Particularly annoying is the stripping of the modern place names, which are the only identifying features for the sites without ancient toponyms.”

### Search functions

It is rather easy to find data. A user only has to fill in the place’s name in the search box on the home page. One can also write only the first letters of a place followed by an asterisk \*. There is also an advanced search function with many optional search criteria.

---

<sup>6</sup> Note on this observation by the user: Pleiades calls itself a “historical gazetteer and graph of ancient places”, thus the observation of the user is in line with what the site itself purports to be.

The research tool is the 'Browse Places' site, which includes a search box, and returns a list and a map of relevant places. The individual results then enable a registered user to propose edits, though the procedure is rather unclear.

LU 2 explains the approach. The basic search function simply returns data for the place name inserted. The Advanced search allows you to search text, titles and descriptions; the latter allows you to search for, viz, 'villas' etc. This is also possible in Categorization, which provides drop-down tables. The first, 'subjects,' is very bizarre, including a number of Roman, Greek and Egyptian gods, some objects, like 'altar', the Latin League, 'palace', 'accuracy', or 'concept' and 'copper'. It is very hard to work out what these words have in common.

The feature categories are more straightforward (amphitheatre, aqueduct etc.) but not exhaustive, lacking in particular production sites, such as kilns, workshops and so on. It does not seem possible to filter a search. This is a particular problem if one wants to filter by a time dimension. There is a vast list of these in vocabs, though, sadly, it does not apply to North Africa.

LU 2 concludes that, basically, the search function is fine for looking up single toponyms, but much less so for more general searches. A slight annoyance was that the site does not return the coordinates to the place without what they refer to as 'drilling down' – downloading the csv table with the lat-long or taking the KML to Google Earth. They suggest it should be possible for the map to return the lat-long, as Google Earth does.

### **Options for downloading data**

Download Data allows the user to download CSV and KML files, though it is unclear how to filter these. It is possible to download a KML file, though this proves so heavy that it is very difficult, if not impossible, to manipulate it on Google Earth.

LU 2 remarks: "If you need a data dump of tables summarizing published locations it is useful, as is the KML dump of the places (except for the difficulty of then using Google Earth with the full KML listing downloaded)."

### **Options for uploading data**

For the registration procedure to join Pleiades, it was suggested that it would be easier to have an empty field where one can enter one's email address, instead of properly emailing to request username and password. A good practice, by contrast, in the exchanges with contributors was seen in the fact that there is a log history available for every record with its history of user inputs. LU 2 attempted to add an excavation website, with little success. "There is a whole file called 'I'd like to add a link to an excavation website – but where does it belong?'"

### **Support services**

The user appreciated the vocabularies section: "It is extremely valuable since many terms and controlled vocabularies are employed by Pleiades including time periods and place categories, which are explained here." There is also a section with scientific presentations and papers about the project and another one with official project reports. These can be easily used as references.


There are a lot of FAQs, with responses, though these were not always found to be very helpful.

## **Conclusions for ARIADNE**

The following suggestions could be taken into consideration in the design of the ARIADNE portal:

- Good practice: the provision of content in the form of maps and coordinates
- Good practice: providing a log history for every record with its history of user inputs
- Learning point: avoid stripping of place names and the lack of reference to non-English resources

### 4.2.8 Portal to the Past

Website	<a href="http://www.portal.cig-icg.gr">http://www.portal.cig-icg.gr</a> 
Provider	The Canadian Institute in Greece
Domain/thematic focus	Archaeological projects of CIG
Geographic focus	Greece
Type of data offered	Digital images and multimedia files, such as aerial photographs, satellite images, maps, GIS images, survey transects and collection points, site plans, trench plans and sections, architectural plans and sections, reconstructions of ancient buildings, drawing of artefacts.
Further information	<p>The “Portal to the Past” highlights the archaeological work of the Canadian Institute in Greece (CIG) since 1980. Currently, the portals contains a representative sample of the imagery and information related to about 20 archaeological research projects carried out by CIG since 1980. The information and imagery is searchable by project, site, find, image, institution, researcher, research expertise and other criteria. The discoveries span from the Mesolithic period (ca 9th millennium BCE) to the 20th century CE. One can search each component for specific information by project, site, images or finds. The languages used in the portal are English, French and Greek.</p> <p>The information and imagery offered by the portal are intended for personal/non-commercial use only. This usage is covered under a Creative Commons License which the Board of Directors of the CIG has adopted for the users of the portal.</p>
Questionnaire ID	20

## Assessment of the portal – strengths, weaknesses, innovative features

The evaluation team chose to conduct research on ceramics as a test for evaluating the portal.

### Overall structure of the portal

The portal is organised quite well. Thanks to the various sections it is easy to navigate in a direct way to find the requested information. With "live link" you can have a quick and easy navigation that takes users from one "layer" of information to another (it will lead the user directly to the particular component, site, find, person, institution), thanks to a hierarchical internal structure. The information is available in three languages: English, French and Greek. Data is available under the CC-BY-NC-ND 3.0 license.

The evaluation team recommends the Chronology section as a good practice for the standardization of information, due to the presence of an internal thesaurus and the availability of detailed metadata of the digital objects.

### Search functions

The evaluators describe their search experience: "Starting a research on ceramics from the excavations in Attica we preceded searching "Attiki pottery" in the search bar in the upper left side. It was not possible to receive information for the above request. The message was: "Your search yielded no results." By placing the query "Attica" we received generic information regarding hyperlinks where the word "Attica" has been indexed. Another way to search for information and pictures of pottery from excavations in Attica is to go to the Sites from where you can see the list of excavations in Attica. By clicking on the link you are directed to the excavation of Kiapha Thiti / Kontra Gkiliate (the only excavation present in that region). Thereupon you get various kinds of information on the excavation: location, map, the site related finds, the site related images. If you click on "Type B4 cooking vessel from Kiapha Thiti (Attica)" you will have information about the digital image, and the "related find"/ "related archive" which have different entry numbers (KTEP011; AO0334) but they represent the same digital object in the others sections (FINDS, ARCHIVES). You can download the image.

The evaluators suggest as a good practice the live links system which is used to easily browse the various sections of the portal. Starting from an object, users can retrieve information about the provenance of data using the hyperlinks.

There is also the possibility to find similar archives/data/information based on query results.

### Mechanisms for downloading data

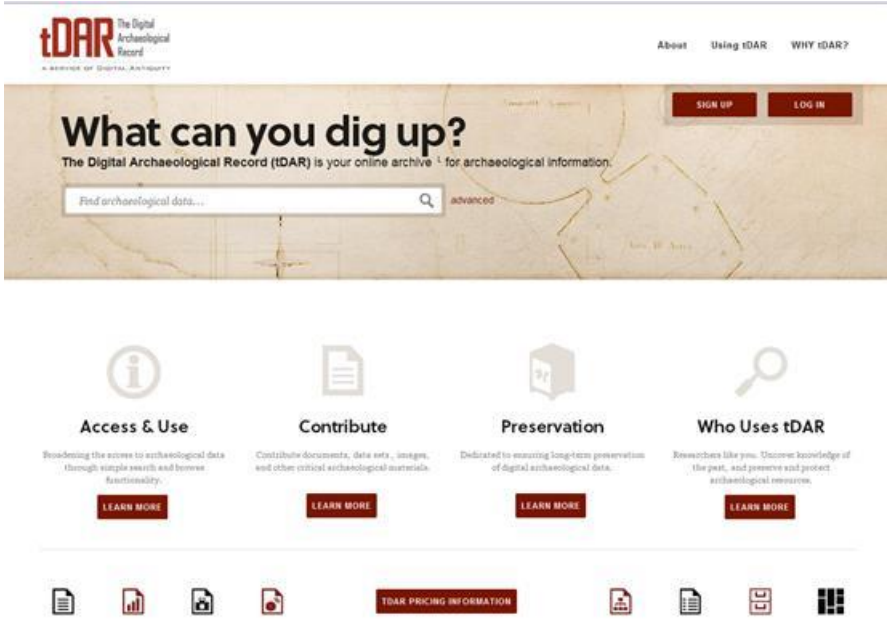
It is possible to download images in JPG/JPEG format.

## Conclusions for ARIADNE

The following suggestions could be taken into consideration in the design of the ARIADNE portal:

- Good practice: the Chronology section for the standardization of information, due to the presence of an internal thesaurus and the availability of detailed metadata of the digital objects.
- Good practice: the live links system which is used to easily browse the various sections of the portal. Starting from an object, users can retrieve information about the provenance of data using the hyperlinks.

## 4.2.9 tDAR – The Digital Archaeological Record

Website	<p><a href="https://www.tdar.org">https://www.tdar.org</a></p> 
Provider	Digital Antiquity, a multi-institutional organisation that has been explicitly designed for the long-term maintenance of tDAR, currently being incubated by Arizona State University
Domain/thematic focus	Digital records of archaeological investigations
Geographic focus	Global
Type of data offered	Different types of data, in particular images, coding sheets, project data, 3D and sensory data, GIS data, ontologies
Further information	<p>The Digital Archaeological Record (tDAR) is a major international digital repository for the digital records of archaeological investigations. tDAR started out as an NSF project and has grown into an international archive for archaeological data. In 2011, The National Archaeological Database (NADB), which captured and catalogued over 350,000 citations for archaeological reports or related materials, was integrated into tDAR, a major milestone in tDAR's development. In 2012, NSF funded refinement of tDAR's data integration interface and a major research application of these tools to large datasets of archaeological fauna from the Southwest.</p> <p>Users do not have to be registered to browse the public metadata in tDAR, but must be registered in order to contribute data to tDAR. Searches can be made in any or all of the fields containing rich metadata associated with each record. Metadata are indexed by major search engines.</p> <p>tDAR offers tools that enable researchers to integrate multiple data sets from different sources and to collaborate with other researchers to create large aggregated data sets.</p>
Questionnaire ID	23



## Assessment of the portal – strengths, weaknesses, innovative features

The evaluator has browsed several tDAR records to see what he would find and how the findings and the functionalities of the portal compared to that of other data archives he was familiar with.

### Overall structure of the portal

It was easy to find and see the relevant information, but there were some downsides.

- The evaluator found some of the lay-out differences on the portal “distracting”. The initial, main search interface is very different in layout (including the “limit by geographic region” map) than options from the search results screen. This gives a disconnected impression.
- Some elements of the portal documentation were not easy/intuitively to find, but “hidden” in different sections. For instance, to reach the FAQ, it was necessary to go to ‘Using tDAR’ => ‘Help & Tutorials’ => ‘Searching tDAR’ (which takes one to a different site) => FAQ. All these pages have a different lay-out. And it is unclear why one is suddenly directed to a different site (‘Confluence’).
- It took more time than expected to start using the Digging/browse page.

### Search functions

The good point was that the options for viewing and filtering search results are very good.

The main weakness is that many results seem not to show up on the map.

### Mechanisms for downloading data

The downloading process seemed to be easy at first sight, but then some challenges occurred.

The user was not sure whether it was possible/easy to download multiple files belonging to the same record. It turned out difficult to find any; most records seemed to have only a single file associated with them. One of the datasets (the EMAP Ceramics Database project) had two excel files, but there was no obvious way how to jointly download both files – this indicates that all files have to be downloaded individually.

Another challenge is that codebooks and reference sheets seem to be only included within the tDAR record metadata. It would be better to include such information with file downloads.

### Options for uploading data

Contributing data involves payment and was therefore not tested. Concluding from the documentation, it seems to be an easy, intuitive process.

## Conclusions for ARIADNE

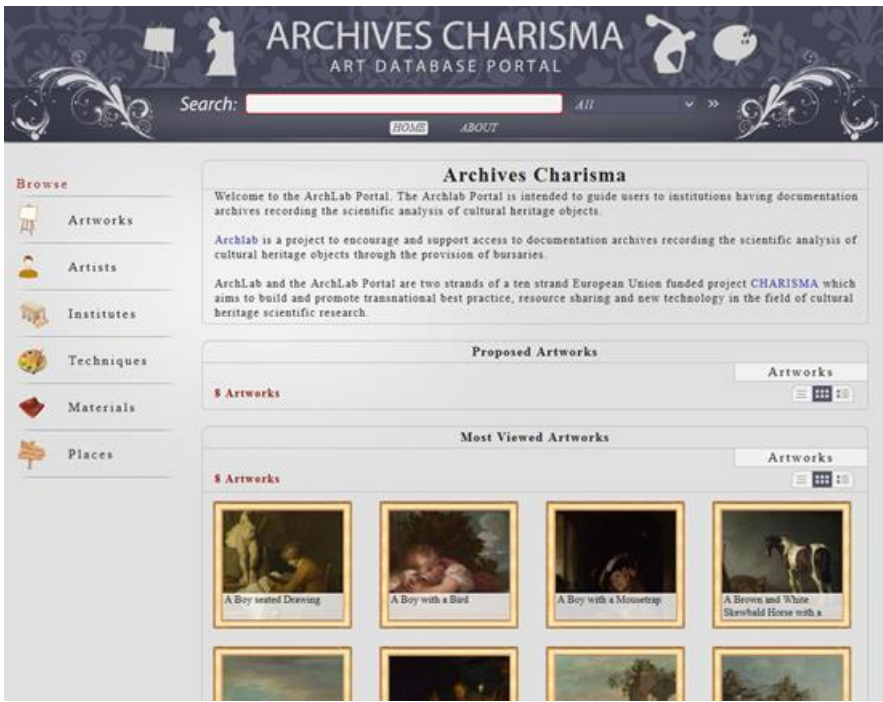
The following suggestions could be taken into consideration in the design of the ARIADNE portal:

- Recommendation: It is important to offer downloading multiple files belonging to the same record in one go, rather than having to download individual files sequentially.



## 4.3 Specialist archaeology portals

### 4.3.1 Archives Charisma – Art Database Portal

Website	<p><a href="http://archives-charisma-portal.eu">http://archives-charisma-portal.eu</a></p> 
Provider	ARCHLAB/CHARISMA project
Domain/thematic focus	Analytical and technical data on cultural heritage objects and artworks
Geographic focus	Europe
Type of data offered	Descriptive data on objects and scientific data (image, spectrum, SEM, XRF, FT-IR, colour measurement)
Further information	<p>The mission of this portal is to strengthen and complement the physical access to the ARCHLAB (the archives of European Museums and Cultural Heritage Institutions) of the CHARISMA Project, a partnership that represents archives of a large and mostly unobtainable number of stored knowledge and technical data held by the most prestigious European museums and conservation institutes from France, UK, Italy, Spain and Netherlands. The portal provides virtual access to a large research community of professional users of such data, via a project for the integration of technical metadata coming from these archives and streamlined via a cultural heritage web portal in a regular, standardised, and consistent way.</p> <p>Data provided by the project partners has been expressed in RDF, configured to the CIDOC-CRM ontology. At present the database shows the name of artist, period, type (e.g. "painting"), institution and place where the object can be found for each object, and information about the colour measurement. For most objects, however, the database does not offer pictures.</p>
Questionnaire ID	07

## Assessment of the portal – strengths, weaknesses, innovative features

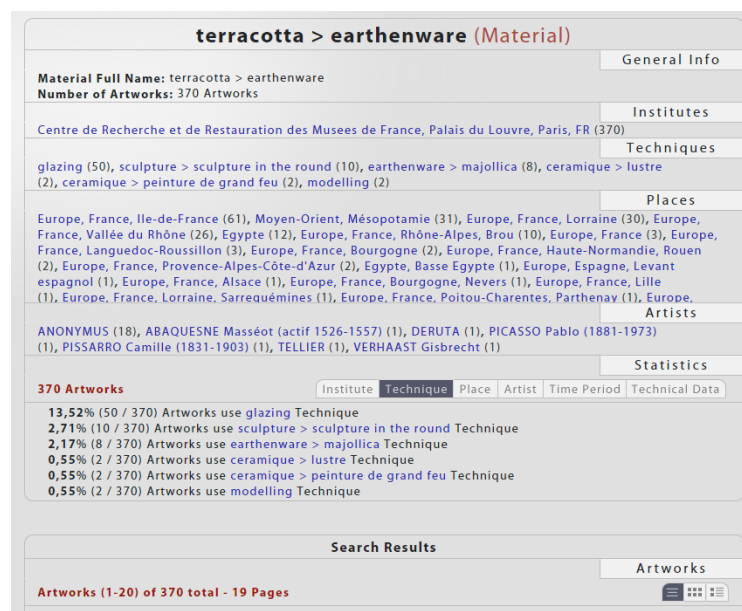
The evaluator specified a hypothetical research task for carrying out the assessment of the portal's functionalities. They wanted to find out scientific data on ceramic objects.

### Overall structure of the portal

The design and structure were found to be logical, and the amount of information to be adequate. However, what was missed was a filtering option for scientific data. One can access scientific metadata (whether xrf results exist on that object or not) through several filtering procedures, which seemed illogical. In particular, the user criticises that an art objects database without pictures was not useful, unless one specifically knows what one was looking for.

### Search functions

The search functions, and in particular the filtering steps and the way filtered data is displayed, were rated as "very good". For example, when the user filtered for "earthenware" (after having selected "materials" on the start page), they got a list of earthenware at the bottom of the page (but without pictures). The system also delivers filtered information on institutes, techniques, places and statistics. On this basis, it was possible to further refine the research within earthenware.



The user suggests, as a further improvement, that it would be useful to explain the way the database is organised by means of some screen shots.

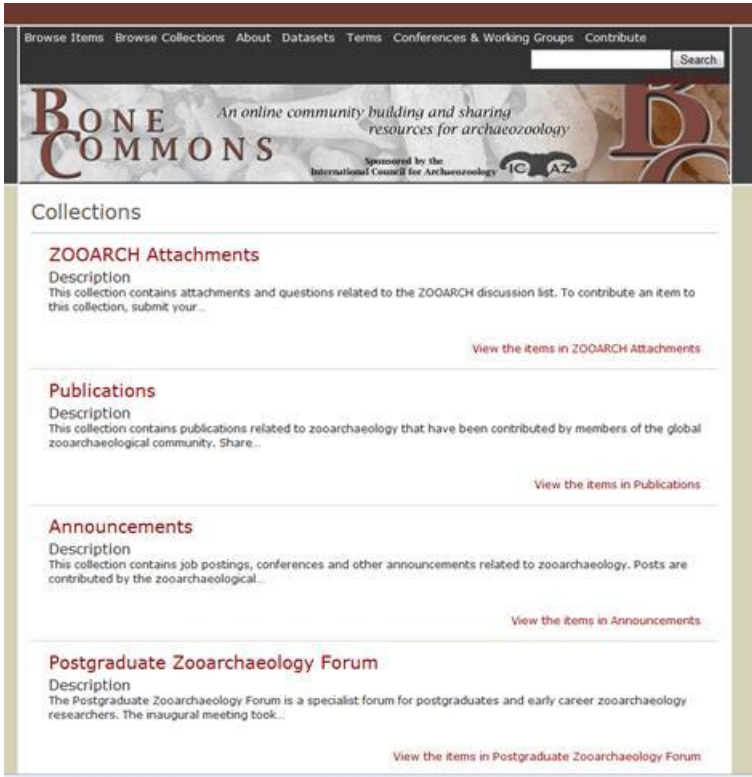
An important drawback was that the database does not say right from the start that it only contains metadata: *"It took me an hour to realise through several filtering and searches that actually there is no data here whatsoever (except metadata), not even pictures."* This can be frustrating for researchers who are not familiar with the database.

## Conclusions for ARIADNE

The following suggestions from the evaluation of Archives Charisma by a lead user could be taken into consideration in the design of the ARIADNE portal:

- Important learning point: Any portal/archive should clearly state what type of data it contains. For instance, if it contains only metadata, this should be clearly stated, to avoid researchers having to spend much time to find out what is "inside".

### 4.3.2 Bone Commons

Website	<a href="http://alexandriaarchive.org/bonecommons/collections">http://alexandriaarchive.org/bonecommons/collections</a> 
Provider	International Council for Archaeozoology (ICAZ), a non-profit organisation devoted to promoting archaeozoological research of the highest scientific standards, and fostering communication among the international community of archaeozoologists.
Domain/thematic focus	Material related to archaeozoology
Geographic focus	Turkey (263,164 objects), Iran (31,970), Germany (20,238), Israel (7024), Jordan (4472), Italy (4225), United Kingdom (2810), United States (2305)
Type of data offered	Mostly osteological (animal fauna) data The search filters are either by spatial context, by project/collection, by category (animal bone/shell), or by classification.
Further information	BoneCommons is an ICAZ-sponsored project developed by the Alexandria Archive Institute. It was launched in May 2006. BoneCommons aims to facilitate discussion and contact between archaeozoologists worldwide by offering forums where ICAZ members can post papers, images, teaching resources, questions and comments. All content on BoneCommons, while owned by the creator of the content, is openly viewable by the general public worldwide. Anyone can search BoneCommons and view its content.
Questionnaire ID	05

## **Assessment of the portal – strengths, weaknesses, innovative features**

The evaluator specified a hypothetical research task for carrying out the assessment of the portal's functionalities. The user wanted to see if there was a correlation anywhere between the research conclusions and primary data used, and if provenance data has been recorded.

### **Overall structure of the portal**

The user rated the overall structure of the portal as "fair": "There are very basic options for advanced/combined searches, no online tools for analysing data and apparently no clear administration of data (some links are broken). In some cases data can be downloaded in CSV format." A good practice was the publication of the ontology in use.

### **Search functions**

The search functions which the portal offers are rated as "poor", as only a keyword search with basic filters (based on set-up terms) of free words search is available.

### **Mechanisms for downloading data**

The mechanisms for downloading data are seen as average. There are several steps needed to arrive in the download section, and some links were apparently broken.

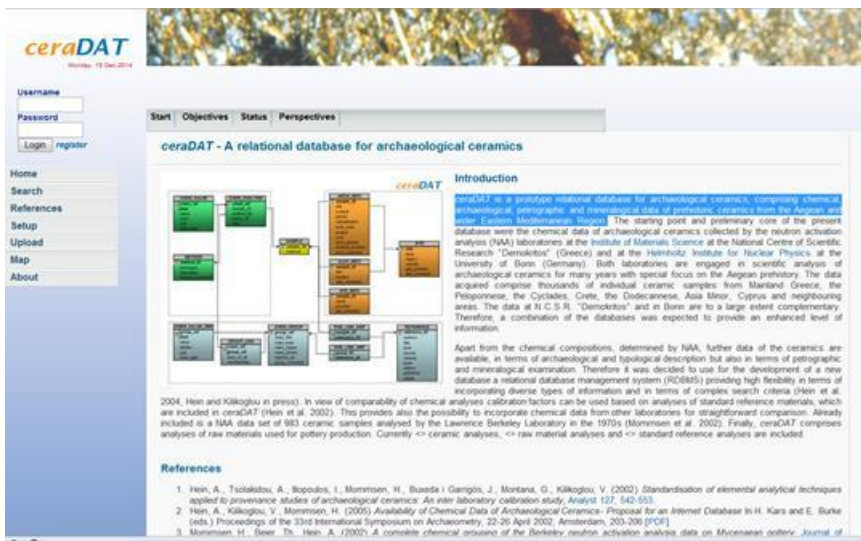
### **Options for uploading data**

The portal provides the opportunity to upload/deposit research data. This function was found to be good – the process proved to be easy and straightforward.

### **Specific support functions & features offered**

There are few specific features offered.

### 4.3.3 ceraDAT

Website	<p><a href="http://www.ims.demokritos.gr/ceradat/?PHPSESSID=dnjxlxkn1">http://www.ims.demokritos.gr/ceradat/?PHPSESSID=dnjxlxkn1</a></p> 
Provider	Institute of Materials Science at the National Centre of Scientific Research "Demokritos" (Greece) and the Helmholtz Institute for Nuclear Physics at the University of Bonn (Germany)
Domain/thematic focus	Archaeological ceramics
Geographic focus	Mediterranean
Type of data offered	ceraDAT is a prototype relational database for archaeological ceramics, comprising chemical, archaeological, petrographic and mineralogical data of prehistoric ceramics from the Aegean and wider Eastern Mediterranean Region. Apart from the chemical compositions, determined by NAA, further data of the ceramics are available, in terms of archaeological and typological description but also in terms of petrographic and mineralogical examination.
Further information	The starting point and preliminary core of the present database were the chemical data of archaeological ceramics collected by the neutron activation analysis (NAA) laboratories at the Institute of Materials Science at the National Centre of Scientific Research "Demokritos" (Greece) and at the Helmholtz Institute for Nuclear Physics at the University of Bonn (Germany).
Questionnaire ID	06



## Assessment of the portal – strengths, weaknesses, innovative features

The evaluator specified a hypothetical research task for carrying out the assessment of the portal's functionalities. The user was looking for a geochemical composition of Neolithic ceramics from Greece.

### Overall structure of the portal

The user rated the overall structure of the portal as “fair”, as it was not found to be designed in a user friendly manner: “The search platform expects you to know the regions and areas of Egypt, Greece, Turkey, Spain etc.” Moreover, NAA data were not collated, for a site but available for each sherd on a separate tab. The user notes that this can turn into a cumbersome procedure: “Imagine when you want to compare 100 sherds from a site – that is a lot of copying and pasting into an excel sheet.”

### Search functions

The user notes that multi-layered drop down menus are provided; however, in order to use them, one “really needs to know the countries and their regions and areas”. The lead user describes in detail the process of his search and his experiences, and documents this with screen shots. The following paragraphs are quotes from the evaluation report.

*One can access this through drop down menus. After selecting Greece in countries you have to select region and area. Here it would be useful to have a map to see the regions and areas, unless you are a Greece expert and you now all the regions and areas. Once you are over these you can select a site. If you don't know regions and areas the easiest way is to filter for Greece and in the period type Neolithic.”*

*The results were not very organised. I received a “text flow” a list of sites. Here, again, a map would have been helpful to see the distribution of these sites within Greece.”*

Archaeology Chemistry Petrography Mineralogy

((site = 'Achaia, various sites' OR site = 'Achinos' OR site = 'Afantou' OR site = 'Aghia Kyriaki' OR site = 'Aghia Photia' OR site = 'Aghia Photia (clay)' OR site = 'Aghia Triada' OR site = 'Agia Irini' OR site = 'Akraiphon' OR site = 'Akropotamos' OR site = 'Akrotiri' OR site = 'Alantors' OR site = 'Alimos' OR site = 'Amfikleia' OR site = 'Angelohori' OR site = 'Ano Moulia' OR site = 'Antikythera' OR site = 'Aqassia' OR site = 'Argos' OR site = 'Armeno' OR site = 'Asine' OR site = 'Atalanti' OR site = 'Athens, Acropolis' OR site = 'Athens, Agora' OR site = 'Attica, various sites' OR site = 'Avli' OR site = 'Berbati' OR site = 'Cap Kolona' OR site = 'Cap Plakakia' OR site = 'Chalki' OR site = 'Chania' OR site = 'Chania (town)' OR site = 'Chomatasi' OR site = 'Chora Ano Englianos' OR site = 'Debia' OR site = 'Dhaskalio' OR site = 'Dimitra' OR site = 'Elefesi' OR site = 'Elefina' OR site = 'Epidauros' OR site = 'Eutresis' OR site = 'Fidokoryi' OR site = 'Frantzi, Larnia' OR site = 'Galepsos' OR site = 'Giantisa' OR site = 'Gla' OR site = 'Hagios Stephanos' OR site = 'Heraklion' OR site = 'Hersonissos' OR site = 'Hill 133' OR site = 'Ialysos' OR site = 'Kalamianos' OR site = 'Kalamonas' OR site = 'Kalapodi' OR site = 'Kalochofatis' OR site = 'Kamiroi' OR site = 'Kanakia' OR site = 'Kardamaina' OR site = 'Kardamaina (clay)' OR site = 'Kastelli' OR site = 'Kastri, Thasos' OR site = 'Katakolo' OR site = 'Kavos' OR site = 'Keramidiaki' OR site = 'Kifissos' OR site = 'Knossos' OR site = 'Kolona' OR site = 'Kommos' OR site = 'Kopreza' OR site = 'Korakou' OR site = 'Koropi' OR site = 'Kos (town)' OR site = 'Kryi' OR site = 'Lake Vouliagmeni' OR site = 'Lamia, Valagouras' OR site = 'Lastros' OR site = 'Lerna' OR site = 'Loudra Elitheron' OR site = 'Makrygiolos' OR site = 'Malis (clay)' OR site = 'Megaplatanos' OR site = 'Mochlos' OR site = 'Mycenae' OR site = 'Myrtos (clay) I' OR site = 'Myrtos (clay) II' OR site = 'Myrtos (clay) III' OR site = 'Myrtos (clay) IV' OR site = 'Myrtos (clay) V' OR site = 'Neo Monastiri' OR site = 'Nopieia' OR site = 'Olympia' OR site = 'Olympia-Kolosakos' OR site = 'Orchomenos' OR site = 'Palakastro' OR site = 'Palakastro (clay) I' OR site = 'Palakastro (clay) II' OR site = 'Panormos' OR site = 'Perikaki' OR site = 'Pellana' OR site = 'Pera Galini' OR site = 'Perati' OR site = 'Peristeria' OR site = 'Perivoli Sperchadios' OR site = 'Petras' OR site = 'Petras (clay)' OR site = 'Phaistos' OR site = 'Philoou' OR site = 'Plisidia' OR site = 'Platanos-Renia' OR site = 'Platyvola Cave' OR site = 'Poros-Katsambas' OR site = 'Psathi' OR site = 'Rhodes (town)' OR site = 'Rhodes, general' OR site = 'Sitagroi' OR site = 'Skopi' OR site = 'Sparta' OR site = 'Stavromenos' OR site = 'Stiri' OR site = 'Syrtria' OR site = 'Tanagra' OR site = 'Thebes' OR site = 'Tholos' OR site = 'Thorkos' OR site = 'Thrasos' OR site = 'Troy Serron' OR site = 'Tragana' OR site = 'Val' OR site = 'Vania' OR site = 'Vasiliki' OR site = 'Vasiliki (clay) I' OR site = 'Vasiliki (clay) II' OR site = 'Vathy' OR site = 'Vathypetro' OR site = 'Vianios II' OR site = 'Villanova (Paradisi)' OR site = 'Voudeni, Amygdalia' OR site = 'Vourvatsi' OR site = 'Vourvoulitis' OR site = 'Vrionisi' OR site = 'Zaros' OR site = 'Zervohori' OR site = 'Zygouris (Neolithic)' ) 187 records found in the database 187 of these records are public.

show 30 rows 1 to 30

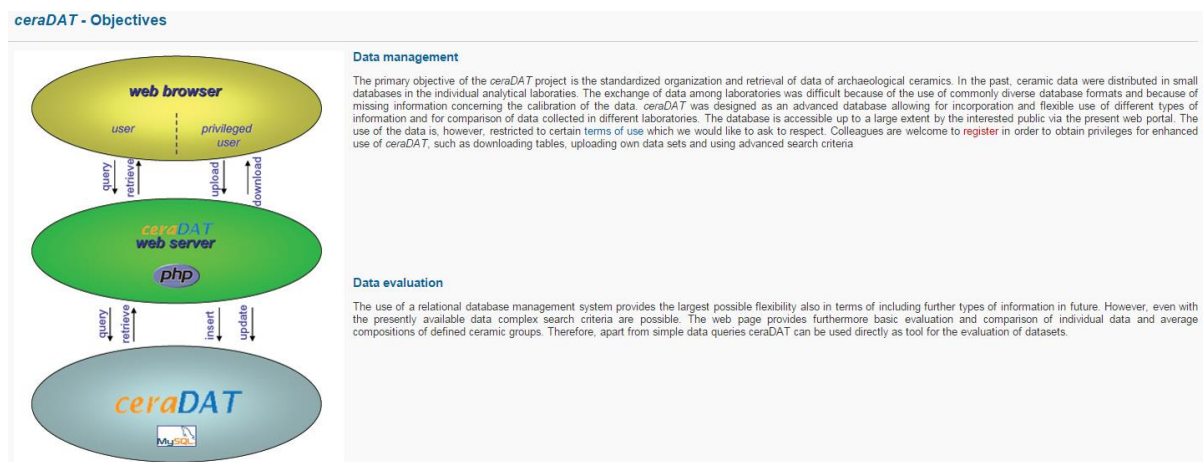
Sample	Site	context	period	classification	arch. code	project	year	partner	location	comment
PLT 2003	Playola Cave		Neolithic	ceramised bowl		PhD Eleni Nodariou	2000	E. Nodariou		Ceramised bowl - NEOLITHIC
AKP1055	Akropotamos	Surface	Late Neolithic	Bowl with 5 profile	AKP1055	Black-on-red pottery	2001	D. Marambaou, Z. Tzafani		Very fine red paste
AKP1059	Akropotamos	Surface	Late Neolithic	Bowl with flared walls	AKP1059	Black-on-red pottery	2001	D. Marambaou, Z. Tzafani		Very fine red paste
AKP1070	Akropotamos	Surface	Late Neolithic	Bowl with flared walls	AKP1070	Black-on-red pottery	2001	D. Marambaou, Z. Tzafani		Very fine red paste
AKP1071	Akropotamos	Surface	Late Neolithic	Bowl with flared walls	AKP1071	Black-on-red pottery	2001	D. Marambaou, Z. Tzafani		Very fine red paste
AKP1072	Akropotamos	Surface	Late Neolithic	Bowl with flared walls	AKP1072	Black-on-red pottery	2001	D. Marambaou, Z. Tzafani		Very fine red paste
AKP1073	Akropotamos	Surface	Late Neolithic	Bowl with flared walls	AKP1073	Black-on-red pottery	2001	D. Marambaou, Z. Tzafani		Very fine red paste
AKP1074	Akropotamos	Surface	Late Neolithic	Two-handled pot	AKP1074	Black-on-red pottery	2001	D. Marambaou, Z. Tzafani		Very fine red paste
AKP1075	Akropotamos	Surface	Late Neolithic	Two-handled pot	AKP1075	Black-on-red pottery	2001	D. Marambaou, Z. Tzafani		Very fine red paste
AKP1076	Akropotamos	Surface	Late Neolithic	Two-handled pot	AKP1076	Black-on-red pottery	2001	D. Marambaou, Z. Tzafani		Very fine red paste

*By clicking a site a short metadata summary appears with a map of that site.*

*By clicking the samples (22 samples in blue) I received a list of samples with their descriptions (site, context, period, classification, project name, year, partner, comment). Here by clicking each sherd NAA data are available for them individually. It would have been useful to make available the collated NAA data for the site to see the compositional variability within that site. Now I have to download each sherd's data and put the tables together manually!*

*I am a registered user, but I could not download tables. I could copy them from the screen. Registered users should be able to download data.*

Here is the access scheme:



### Mechanisms for downloading data

The mechanisms for downloading data are seen as poor, for the reasons explained above – there was no possibility to download coherent data sets.

### Options for uploading data

The upload section was found to be good. It seemed well-organised and easy to use. One can upload NAA data in different data formats. However, an opportunity to upload pictures for the samples seemed to be missing.

### Specific support functions & features offered

There were few specific support functions offered.

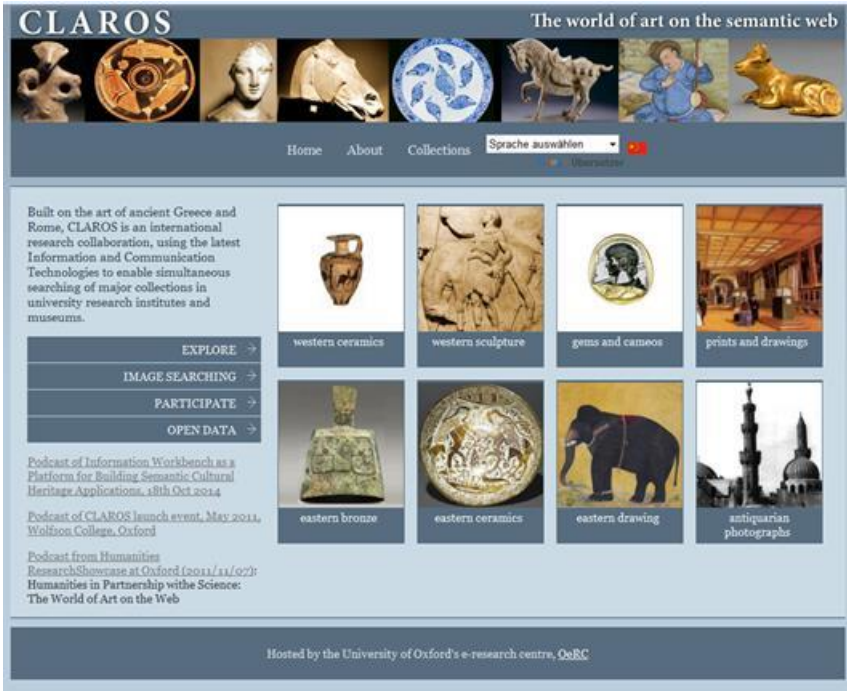
### Conclusions for ARIADNE

The following suggestions from the evaluation of ceraDAT by a lead user could be taken into consideration in the design of the ARIADNE portal:

- Important feature: When users complete a search, they expect that the *collated* NAA data for the site are made available so that they can see the compositional variability within that site, rather providing the NAA data for each item/sample individually.



#### 4.3.4 CLAROS - Cassical Art Research Online Research Services

Website	<p><a href="http://www.clarosnet.org/XDB/ASP/clarosHome/index.html">http://www.clarosnet.org/XDB/ASP/clarosHome/index.html</a></p> 
Provider	University of Oxford's e-research centre (OeRC)
Domain/thematic focus	Classical art and architecture
Geographic focus	Europe, Asia
Type of data offered	CLAROS is a resource discovery service. Its goals are to provide caching, indexing, querying and visualization services, therefore, the data available from the portal is short metadata records associated with online, low-resolution images and links to their originating source.
Further information	<p>CLAROS is an international interdisciplinary research federation, enabling simultaneous searching of major collections in university research institutes and museums. CLAROS began in 2000 as CLassical Art Research Online Research Services with Europe's leading research centres for the art of ancient Greece and Rome. Since 1979 these centres, in the universities of Oxford, Cologne, and Paris, have been creating scholarly databases about the art of classical antiquity.</p> <p>Since 2007 CLAROS has been using semantic web technologies to make the geographically separate scholarly datasets interoperable. While each content provider retains their data, formats and website, the user can search all datasets as one by text or by image. By 2010, more than 20 million records had been made 'interoperable' using the CIDOC-CRM ontology.</p>
Questionnaire ID	09

## **Assessment of the portal – strengths, weaknesses, innovative features**

The evaluator specified a hypothetical research task to carry out the assessment of the portal's functionalities consisting of:

- searching for images relating to the research topic 'Byzantine sculpture' in order to illustrate explanations of sculptural types, and
- uploading images of sculptural forms to discover if the collection has examples of similar forms in order to gain an idea of where similar sculpture is found.

### **Overall structure of the portal**

The user rated the overall structure of the portal as good. The design and structure of the portal were considered fairly logical and easy to navigate. However, overall, the text across the whole portal was seen to be quite small and should be larger to improve readability. The multiple box design structure was found to be complicated and very busy: "These boxes are often also quite small and close together. This could be redesigned to improve visual appeal of portal and prevent people from being put off using the tool by its perceived complicated design."

The user highlights a good practice in the structure: "The option for being able to have the results list at one side of the page and a particular record on the other is good. This makes it easy to move through a long list to find the records you are interested in without having to use back and forward buttons or open new window, as is the save record function."

### **Search functions**

The search functions which the portal offers are rated as average. The user provides reasons for the assessment: "The facets search options are comprehensive and the search results appear quickly, but the text searches have some limitations and reliability issues. When directly accessing the portal it opens in English. But when you use the search options it doesn't recognise English place names. i.e. a search for Turkey returns no results but a search for Türkiye does. When the portal is set to a different language and then reset back to English this issue seems to resolve itself about one in three times.

Frequent use of the text search boxes and multiple facets searches also seem to break the portal.

The user suggests that Claros could be a model in that respect: "I actually found the 'Claros data' (<http://data.clarosnet.org/objects/>) pages designed to accompany the LOD an easier tool to browse as they didn't break."

The ability to carry out a facet type browse was seen as a good practice in the design.

### **Mechanisms for downloading data**

The mechanisms for downloading data are seen as poor, as there were apparently no download options for the results of the specified search. The user suggests: "I would have liked to have seen the ability to download a csv file of any saved searches."

### **Options for uploading data**

Claros provides options for uploading data, but this was seen as fair. The portal offers the opportunity to upload a photo and to search for similar photos in the collection. These images do not enter a collection or archive. The user notes: "I could not get results for similar images despite uploading numerous images that were, to me, very similar to items in the collection. i.e. a bust of Hadrian almost identical to an image in collection. The only time I could get it to work was when I uploaded an image already in the collection."

The actual upload process was found to be "very easy and quick", however.

The user concludes from this experience: "The image upload and recognition tool is a very good research tool, but it needs to work effectively to be of use."

### **Specific support functions & features offered**

The user noted the following support functions that are offered by CLAROS:

- Community services: guidance is available on how to become part of the CLAROS community to share resources
- Communication tools: users are able to add comments to records.
- Information material: The 'about' section of the portal has a very good technological overview and a technical wiki is available for data providers.

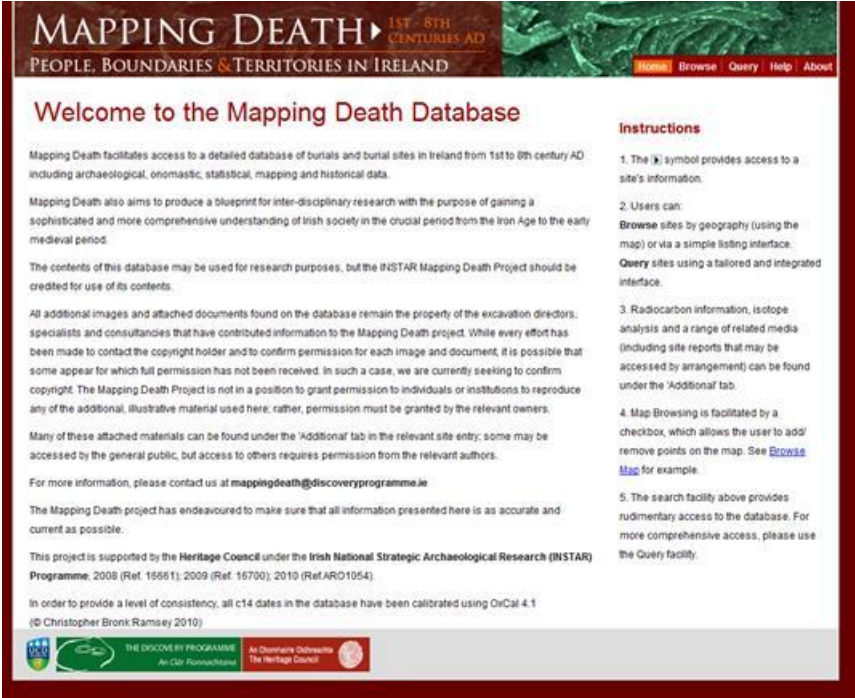
The user remarks that the "help" function is useful; however, while there is a contact email address, there is no active community exchanging views, comments or help with each other. It is suggested that effective help and contact details for further help are essential, and that a section on "Frequently Asked Questions" (FAQ) can be useful.

### **Conclusions for ARIADNE**

The following suggestions from the evaluation of CLAROS by a lead user could be taken into consideration in the design of the ARIADNE portal:

- Best practice: an option for being able to have the results list (of a search carried out) at one side of the web page and a particular record on the other. This makes it easy to move through a long list to find the records the user is interested in (without having to use back and forward buttons or open new window, as is the save record function).
- Good practice and useful feature: the ability to carry out a facet type browse
- Useful features: an image upload and recognition tool (if working well) is a very useful research tool
- Useful features: a contact point for help and a FAQ section are seen as useful

### 4.3.5 Mapping Death

Website	<p><a href="http://www.mappingdeathdb.ie">http://www.mappingdeathdb.ie</a></p> 
Provider	The Discovery Programme
Domain/thematic focus	Burial sites in Ireland from the 1st to the 8th century AD
Geographic focus	Ireland
Type of data offered	Archaeological, onomastic, statistical, mapping and historical data about burial sites
Further information	Mapping Death aims to facilitate access to a detailed database of burials and burial sites in Ireland from 1st to 8th century AD. The database aims to produce a blueprint for inter-disciplinary research with the purpose of gaining a sophisticated and more comprehensive understanding of Irish society in the crucial period from the Iron Age to the early medieval period.
Questionnaire ID	14

## **Assessment of the portal – strengths, weaknesses, innovative features**

### **Overall structure of the portal**

The overall structure is clear and makes it easy to understand how to find the relevant information and data. The boxes with the descriptions were found to be too small, however, and headlines are not clearly emphasized. The user recommends that more pathological data should be entered into the database. Also, it would be useful to add data for metrical measurements.

The user found it useful that there are short interpretations of the sites including both the archaeological and the osteological data.

### **Search functions**

All the options and types of information (such as the site or period) which are available can be searched individually in the database. The user said it was good that it is possible to filter by certain archaeological dimensions (such as periods) or by osteological data (female/male burials or particular traumas). They also appreciated having a map available for browsing between sites.

### **Mechanisms for downloading data**

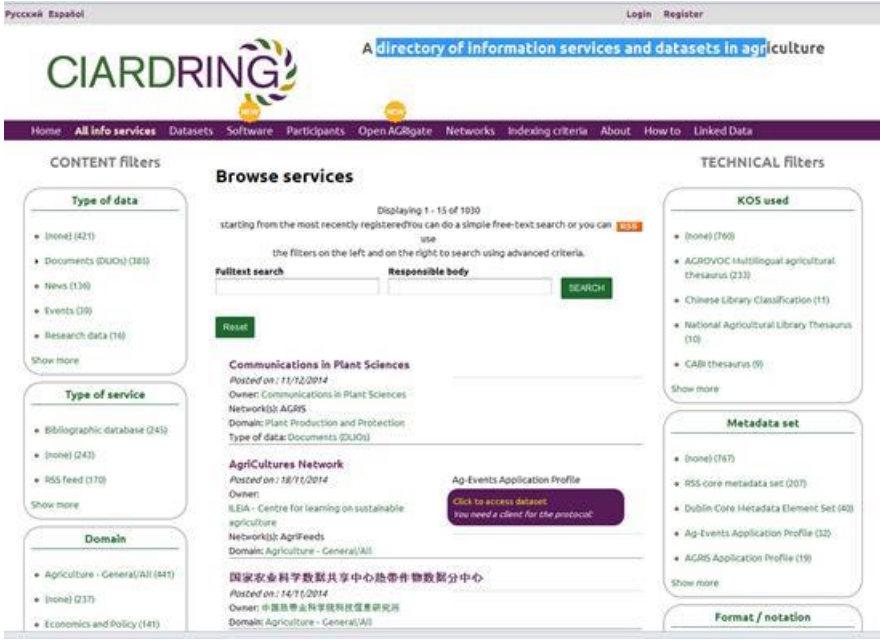
There is no possibility for downloading or uploading data.

## **Conclusions for ARIADNE**

The analysis of this portal has unfortunately not delivered any good practices or recommendations for ARIADNE.

## 4.4 Portals in other domains

### 4.4.1 CIARD Routemap to Information Nodes and Gateways (RING)

Website	<p><a href="http://ring.ciard.net/">http://ring.ciard.net/</a></p> 
Provider	Global Forum on Agricultural Research
Domain/thematic focus	Agriculture (in particular agricultural research for development - ARD)
Geographic focus	Global
Type of data offered	CIARD RING provides web-based information services and datasets about agricultural research for development.
Further information	<p>CIARD RING is a project implemented within the Coherence in Information for Agricultural Research for Development (CIARD) initiative and is led by the Global Forum on Agricultural Research (GFAR). The RING aims to provide an infrastructure to improve the accessibility of the outputs of agricultural research and of information relevant to agricultural research for development (ARD). The CIARD partners intend that the RING will become the principal global technical platform for accessing, sharing and exchanging datasets. The registry was launched in 2009. In 2013 the initiative started its second phase, under the agINFRA project. It aims to leverage the metadata about dataset interoperability in the RING to support other information systems automatically.</p>
Questionnaire ID	08

## Assessment of the portal – strengths, weaknesses, innovative features

The evaluator of this portal focused on checking and assessing the overall functions, rather than carrying out specific search, as they were not familiar with the specific theme and content of that portal.

### Overall structure of the portal

The user rated the overall structure of the portal as “fair”. The portal is a rather simply designed site, with a few, easy to understand pages and sections. On the start page, the main sections appear as headers, and there are tables that immediately inform the user on new services provided, new datasets, providers etc. This gives a useful overall picture of the portal. The user appreciated the provision of a news feed on the home page.

### Search functions

The search functions which the portal offers were found to be good, even if the main search pages could be more evident, instead of just stating “All info services” and “Datasets”. The user noted as a good practice the division of filters available into “Content Filters” and “Technology Filters”. These filters are present in the two basic search pages, “Browse Info” or “Datasets”. Thus, the user can easily set the filters he needs for his research.

### Mechanisms for downloading data

The mechanisms for downloading data are seen as “good”, as the download section was very clearly marked on each page that presented the information or dataset concerned.

### Options for uploading data

The portal lets users upload data. The sharing mechanism was found to be very well organised. It guides the user through tabs to the steps and information needed, and is available in five languages. It has clearly defined units and seems easy to use. (The user did not actually upload data, however.) The tabs that guide the user to uploading and documenting the data that they are going to share were recommended as a good practice. The user found that this was much preferred when compared to a simple continuous page with headers.

### Specific support functions & features offered

The site offers the following features:

- A news feed regarding the new material available in the portal (on the home page), giving an overall picture of the portal's material.
- Indexing criteria: information regarding the standards that the RING uses to index information services, which is very helpful in order to understand the science and technologies related to the portal at a glance.
- A help page called “How To”, which has detailed instructions in various themes as web pages or in pdf format.

## Conclusions for ARIADNE

The following suggestions from the evaluation of CIARD by a lead user could be taken into consideration in the design of the ARIADNE portal:

- Good practice: the division of search filters available into “Content Filters” and “Technology Filters”
- Good practice: tabs that guide users for uploading and documenting the data that they are going to share through a portal.



## 4.4.2 Eurostat

Website	<a href="http://ec.europa.eu/eurostat">http://ec.europa.eu/eurostat</a> 
Provider	Eurostat, the statistical office of the European Union
Domain/thematic focus	Official statistics of the European Union
Geographic focus	Europe (EU)
Type of data offered	Official statistical data (mostly in tables that can be downloaded) on industry, finance, population, trade, transport, environment, energy, science and technology
Further information	<p>Eurostat, established in 1953, is the statistical office of the European Union situated in Luxembourg. Its task is to provide the European Union with statistics at European level that enable comparisons between countries and regions. Eurostat collects and provides a whole range of data that governments, businesses, the education sector, journalists and the public can use for their work and daily life.</p> <p>Eurostat has a policy of encouraging free re-use of its data, both for non-commercial and commercial purposes. All statistical data, metadata, content of web pages or other dissemination tools, official publications and other documents published on its website (with few exceptions) can be reused without any payment or written licence, provided the source is acknowledged, and when re-use involves modifications, this must be stated clearly to the end user.</p> <p>The Eurostat portal offers statistics ("data") as well as publications where data are commented on and put into context. Data stemming from surveys are mostly available as descriptive statistics (in aggregate); the micro (case-level) data are typically not available for download.</p>
Questionnaire ID	25

## Assessment of the portal – strengths, weaknesses, innovative features

The evaluator specified a hypothetical research task for carrying out the assessment of the portal's functionalities. This task consisted of searching and downloading structural statistics for specific industrial sectors (automotive industry, chemical industry) and for reports on these industries.

### Overall structure of the portal

The overall structure of the new Eurostat website was seen as "very good". It does not only have a modern and attractive design, but also provides directly at the start page different quite intuitive routes for searching. There are two main selections on the start page:

- In the top navigation bar, the user can choose between "data" and "publications" (apart from the "news" and "about" section). Here, data means tables with figures, while "publications" feature data in a commented and contextualised way.
- On the start page, the user can search by topic. Nine topics are offered.
- There are further entry points to the rich content of the Eurostat portal in different blocks. One of them (on the top right corner) is named data. Here, there is a direct access to "complete databases".

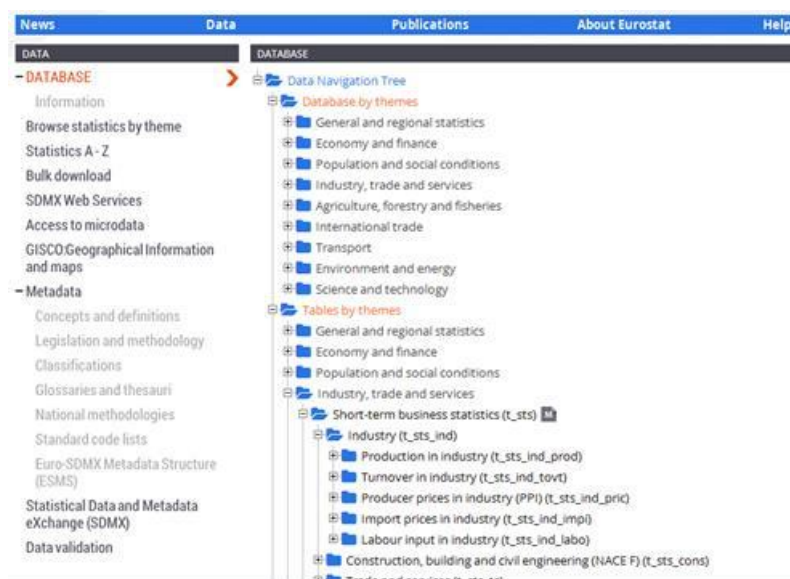
### Search functions

The search functions are extremely rich and diverse, but also confusing. The same content can be reached through different paths, depending on the selection of the main entry gate as described above. For instance, a search for statistics on industry sectors can be started by theme ("Industry, trade and services"), followed by the selection "Structural business statistics"; or it can be started by the type of data searched (selection of "complete database" in the data block), followed by the respective selection in the "data tree" (→ tables by themes → industry, trade and services → structural business statistics). This may appear to be confusing in the beginning, but will lead a user to the same source of data with a few clicks, irrespective of the entry point chosen.

Whether the multiple pathways for finding data and publications are good practices or rather confusing is probably a matter of preference.

The different datasets that are offered by Eurostat are presented in a navigation tree, displaying the datasets in a hierarchical structure (by clicking on a category, the next level opens – see Figure). This is probably a convenient way of organising large data collections.

Figure: Eurostat portal – the data navigation tree



Due to the enormous number of datasets and publications available on the portal, a user who is not familiar with these collections may be overwhelmed and find it difficult to identify the relevant data sets. For researchers who are familiar with Eurostat and their collections, it should be easy to navigate.

### **Mechanisms for downloading data**

The mechanisms for downloading data are quite convenient. Data are offered in different levels of detail and different formats, depending on the data set. Typically, data files can be downloaded and saved as files in xls, html, xml, pdf or csv format. In addition, datasets are available in different levels of detail. For example, the xls dataset from the Structural Business Statistics (sbs) collection “Annual detailed enterprise statistics for industry (NACE Rev. 2, B-E) (sbs\_na\_ind\_r2)” is available in the following options:

- With Footnotes
  - XLS with short description (21.0 KB)
  - XLS without short description (21.0 KB)
- Without Footnotes
  - XLS with short description (14.0 KB)
  - XLS without short description (14.0 KB)

The data sets can be opened and looked at before they are downloaded in interactive viewers. This would theoretically be a very convenient function; but it comes with a challenge: with some of the larger data sets, it can take very long to open them online, so that the procedure gets cumbersome.

The main challenge when working with Eurostat data sets is the many gaps in the data. When downloading recent industry statistics, for example, there is often only information for specific countries. Or, in time series, there are many gaps for specific years. This is a known challenge for researchers who work with Eurostat economic and social statistics.

There is comprehensive information available about the metadata of the datasets, including information on Euro-SDMX Metadata Structure (ESMS), classifications, concepts and definitions, national methodologies, and standard code lists.

All datasets are well explained with regard to what exactly they cover, but sometimes it takes a bit of time to identify the relevant metadata information.

### **Specific support functions & features offered**

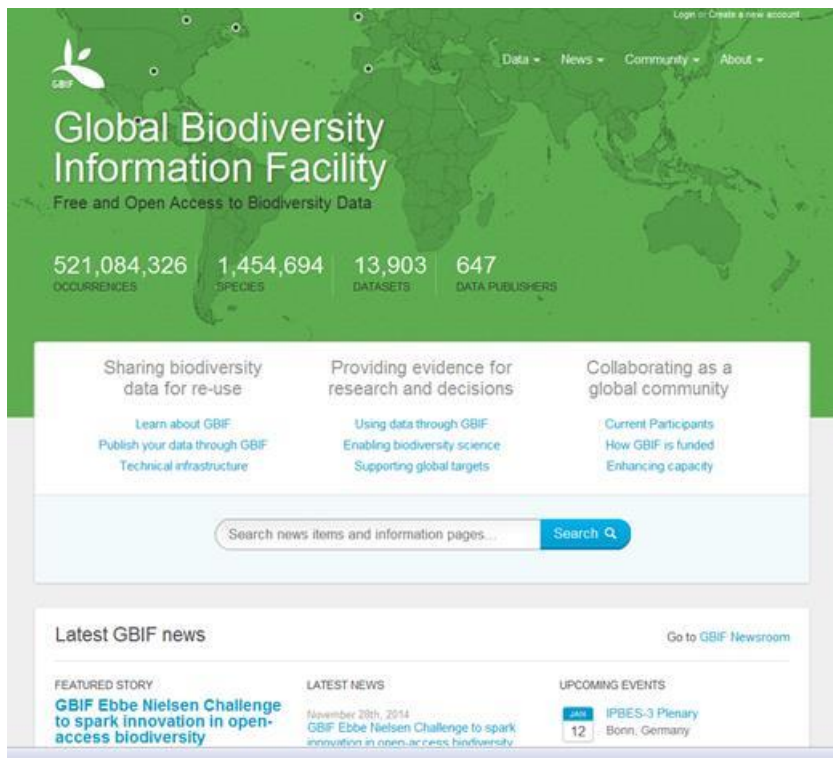
The features are as comprehensive as the data collections, including the opportunity to register and then receive newsletters and information about new publications in selected domains. There is a whole “help” section with different user support mechanisms. Of course the portal also offers a news section, along with, RSS feed functionality, and various other tools.

### **Conclusions for ARIADNE**

The following suggestions from the evaluation of Eurostat by a lead user could be taken into consideration in the design of the ARIADNE portal:

- Good practice: offer different routes to find the (same) relevant data sets, for example through searching by topic or by type of data
- To be considered: whether the main dichotomy between “data” (here: statistics, tables) and “publications” in presenting the materials may also make sense for a portal on Archaeological research

#### 4.4.3 GBIF – The Global Biodiversity Information Facility

Website	<a href="http://www.gbif.org">http://www.gbif.org</a> 
Provider	GBIF (an international open data infrastructure)
Domain/thematic focus	Biodiversity (species)
Geographic focus	Global
Type of data offered	Metadata, occurrences (observations, specimens etc.) and checklists (names). It provides a single point of access to more than 500 million records, shared freely by hundreds of institutions worldwide, making it the biggest biodiversity database on the Internet.
Further information	<p>GBIF is an international open data infrastructure, funded by governments. It was officially established in 2001.</p> <p>GBIF allows anyone, anywhere to access data about all types of life on Earth, shared across national boundaries via the Internet. The data accessible through GBIF relate to evidence about more than 1.4 million species, collected over three centuries of natural history exploration and including current observations from citizen scientists, researchers and automated monitoring programmes. More than 1000 peer-reviewed research publications have cited GBIF as a source of data, in studies spanning the impacts of climate change, the spread of pests and diseases, priority areas for conservation and food security.</p> <p>For researchers, GBIF offers free access and unlimited downloads for all records published via their network.</p>
Questionnaire ID	24

### **Assessment of the portal – strengths, weaknesses, innovative features**

A brief evaluation of the main functionalities and the usability of the site was carried out. The main conclusions from this test are:

- The portal offers a user-friendly environment, elegant design, easy navigation, and logical structure. Good practices are the imaginative underlying interactive map, the easy navigation, and the clarity of the structure.
- Easily accessible "search" on main page, data retrieval is easy and intuitive.
- There are several options for searching. One can just enter a search term - this search result only covers the text content of the news and information pages of the GBIF portal. If one wants to retrieve actual data, it is possible to search by "Publishers and datasets", "Countries", "Occurrences" or "Species".
- Downloading is simple, access is easy. It is necessary to be registered, however. Through the portal, one can easily download records that are published through the GBIF network.
- Uploading and sharing of data is relatively simple. It works in three steps. First time, one has to become a registered GBIF data publisher, then review the GBIF data publishing manuals and select the tools, and finally one has to prepare the data (formats, metadata etc.) and register them with GBIF.
- GBIF offers several free to use, open-source tools and services. They span several categories of use, such as data assessment, data cleaning, data publishing, data visualization, and metadata authoring. A list of tools can be found on <http://tools.gbif.org/>.
- There is apparently no active user community exchanging views on the main pages.

### **Conclusions for ARIADNE**

The analysis of this portal has not delivered any good practices or immediate recommendations for ARIADNE.

#### 4.4.4 GCMD – the Global Change Master Directory

Website	<a href="http://gcmd.nasa.gov">http://gcmd.nasa.gov</a> 
Provider	NASA
Domain/thematic focus	Earth Science Data (including themes such as oceans, climate indicators, sun-earth interactions, atmosphere, cryosphere, paleoclimate)
Geographic focus	Global, but with a strong North American focus
Type of data offered	The GCMD holds more than 34,000 Earth science datasets and service descriptions, using the Directory Interchange Format (DIF). The collections cover subject areas within the Earth and environmental sciences, including land surface data (for landscape analysis) and data by sensor type (e.g. lidar, multispectral satellite data).
Further information	The GCMD evolved from the prototype NASA Master Directory (NMD, first released in 1987) as part of the National Space Science Data Center (NSSDC). In 1994, the GCMD project became part of the Global Change Data Center within the Earth Sciences Directorate at NASA/GSFC, where it still resides. The mission of the GCMD project is to assist researchers, policy makers, and the public in the discovery of and access to data, related services, and ancillary information (which includes descriptions of instruments and platforms) relevant to global change and Earth science research. Searches on the portal can be made by science keywords, instruments, platforms, locations, providers, projects, maps/dates and as free text searches.
Questionnaire ID	12



## **Assessment of the portal – strengths, weaknesses, innovative features**

The two evaluators that focused on this portal specified the following hypothetical research tasks for carrying out the assessment: Lead User 1 imagined they were participating on a research project looking at coastal erosion and wished to find relevant data, services and tools; Lead User 2 tried to find a tool or service to extract metadata from a variety of file formats used for geophysical data.

### **Overall structure of the portal**

Interestingly, the two users (LU 1, LU 2) had different views on the structure – while LU 1 felt overwhelmed by the many different menus on one page, LU 2 appreciated having different search options set out this way and felt their number was “well-balanced” (see below).

LU 1 rated the overall structure of the portal as “average”. They remarked that there was no explanation on entry on what the portal is supposed to do and who it is for. They felt that the design of the interface looked dated and does not cater for any responsive adjustments. Pages and search results are reported to be quite slow in loading. Also, there is no explanation on what each area of the portal does – therefore, it is confusing to early users.

The navigation components were found to be quite text heavy. The user recommends that the development of an icon set for each of them would be better, as images are so small that it is hard to make out what they are.

The user criticizes that there are too many menus within one page. The portal therefore seems overwhelmingly large. This can make it difficult for new users to find the data they require in this or the other portals it cross searches.

LU 2, by contrast, felt the portal was well structured. The primary divisions are between datasets and services/tools (a third, ancillary descriptions, appears to be redundant). Free text search of datasets or tools is provided in a number of places. Users can also browse through the available resources based on science keywords and several other useful entry-point categories.

They even suggested that the presentation of structured access to datasets and services in this portal was a good practice, as users can browse and sort by keywords, instruments, platforms, locations, providers, projects, map/date (though map/date does not appear to be working). The function to sort by different approaches to data in the field accommodates many different types of user and user need. The number of options is well-balanced; too many and the user resorts to free text search, too few and the breadth of datasets and services included would not be apparent from minimal interaction with the tool.

### **Search functions**

LU 1 is not satisfied with the search functions the portal offers. He recommends that, when searching the first page, users should be presented with a Google style free text search rather than having to drill down through all the themes and subthemes. He also suggests that search refinement by location should be nearer the start of the search project for such a global dataset, and that search results should display thumbnail images of the data or representative images of the data to make it easier to distinguish the relevant data sets. All in all, there are too many themes and sub themes presented to users as the main path of data discovery. The time spent exploring down and across these themes can be quite long.

LU 2 regards the search functions as average. The browse function described above serves the user well in navigating through a very large set of resources efficiently. The search function is a single free text box with a radio button choice to search datasets or services. Some users may prefer to make use of the categorisations in the portal through an advanced search. The free text search provided useful results for the dummy task and is optimal for speed and convenience.



They highlight that the portal provides a "Search Tips" page which aligns the behaviour of the tool with that of commercial search engines: matching the expectations of users for 'standard' behaviour of free text search is a useful feature which minimises frustration.

### **Mechanisms for downloading data**

The portal provides a description page for each dataset or service which points the user to an external site for download. Generally, the layout of this page makes it clear to the user where to find more information and where to obtain the data/tool directly, through the use of hyperlinked GET SERVICE/GET DATA and VIEW RELATED INFORMATION. LU 2 argues that this notation speeds up the acquisition process without obscuring information users may require. Other information such as citation, abstracts, keywords, coverage and access constraints provides information that may be required to use the dataset. The user suggests that the presentation of this could be revised, as the the column type display results in scrolling and unused white space on the right of the display.

LU 1 remarks that there is no formal layout or mechanism for this. For users, it can be hard to work out what data exists behind the metadata for some datasets.

### **Options for uploading data**

Users have to register through the EOSDIS User registration system to be able to add data to the portal. Other ways of participating and sharing information with the community include forums. Both evaluators appreciate that there are a number of guidance documents for adding metadata and descriptions on the 'collaborate' page, which would be useful for new users or users adding data under new headings for the first time. They stress that it is good to have guidance for all potential contributors on the areas of metadata creation, licensing and standards. Unfortunately, they are not all clearly laid out within this portal.

### **Value added services**

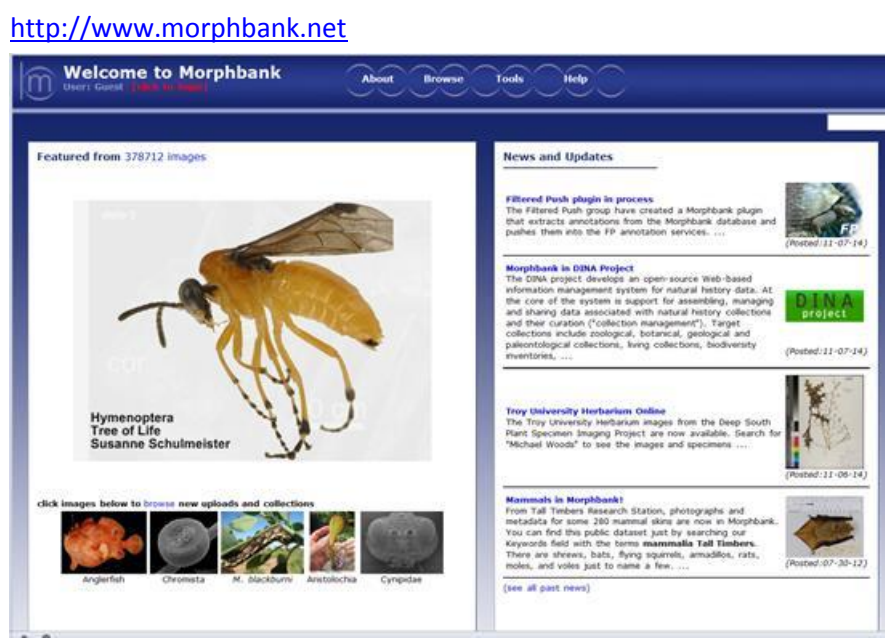
There is a menu option to collaborate through web services alongside links describing various networks of which the GCMD is a part, as well as a 'how to collaborate' option. Community services are accessed through contact and creation of a user account. Documentation is provided on web, keyword and catalogue systems. User interaction is also available on the INTEROP Forum, a mailing list. However, given that the last messages were posted in June 2014 and the archive displays to 2008 without scrolling, vivid exchange of views is not suggested. The latest newsletter available is 2013, with archives to 1997. The docBUILDER tool to add descriptions to the GCMD suggests itself as the most useful support tool for users.

## **Conclusions for ARIADNE**

The following suggestions could be taken into consideration in the design of the ARIADNE portal:

- Good practice: Search results should display thumbnail images of the data or representative images of the data to make it easier to distinguish the relevant data sets.
- Good practice: When searching the first page, users should be presented with a Google style free text search rather than having to drill down through all the themes and subthemes.
- Good practice: Research portals should provide guidance for all potential contributors on the areas of metadata creation, licensing and standards.

#### 4.4.5 Morphbank: Biological Imaging

Website	<p><a href="http://www.morphbank.net">http://www.morphbank.net</a></p> 
Provider	School of Computational Science (SCS), USA
Domain/thematic focus	Specimen-based research in comparative anatomy, morphological phylogenetics, taxonomy and related fields focused on increasing the knowledge about biodiversity
Geographic focus	Global
Type of data offered	Images, text (specimen, taxa, localities, collections, publications)
Further information	<p>Morphbank was established in 1998 by a Swedish-Spanish-American group of entomologists and is currently housed at the School of Computational Science (SCS) at Florida State University. It is a continuously growing database of images, currently consisting of more than 216,000 public images of more than 4500 different species. There are additional images in the system that are presently not public, but are held privately until the contributing scientists are ready to release them. The database aims to facilitate research efforts (and education) by making it possible to store, discuss and share detailed images of specimens from all over the world.</p> <p>The software used in the current system includes PHP, ImageMagick, MySQL, Apache, Java, and JavaScript.</p>
Questionnaire ID	15

## **Assessment of the portal – strengths, weaknesses, innovative features**

The evaluation of the portal's functionalities consisted of systematically trying all features, and assessing them based on a comparison with other databases and search engines known to the tester.

### **Overall structure of the portal**

The user rated the overall structure of the portal as average. The user felt that the features of the page looked nice when visited the first time, but "after that it's mostly an annoyance". For regular use, the user recommends a more prominent search interface already on the home page.

### **Search functions**

The user had some critical comments on how the search is organised on this portal:

- "Sort and filter (limit) are useless, as one can only sort by metadata, not domain content."
- "The search is exceedingly slow."
- "The search term is not highlighted in the results."
- "The search functionality is scattered; some functionalities are even hidden under the "Help" menu."
- "Cannot share direct links to search results."

He suggested that an option to sort by relevance would be useful.

The user also remarked that the buttons relating to the search results are quite small, and their functionality is sometimes opaque. At other times, the functionality was found to be needlessly duplicated, leading to more confusion about where exactly one should click.

### **Mechanisms for downloading data**

The user found download options to be inconsistent. Some searches offer download options (e.g. a csv export), while others do not.

### **Specific support functions and features offered**

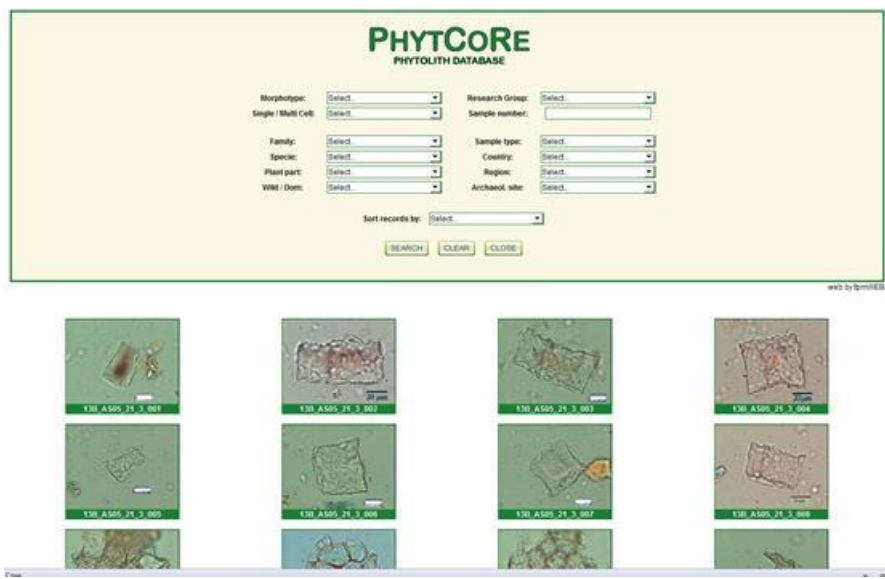
The user remarked that annotations show some promise, but for the time being do not seem to be much used. They were also poorly integrated with the rest of the site (for instance, there was no notification system when items of interest get annotated).

Bookmarking items (collections) was found to be useful.

## **Conclusions for ARIADNE**

The analysis of this portal has unfortunately not delivered any good practices or recommendations for ARIADNE.

#### 4.4.6 PhytCore phytolith database

Website	<a href="http://gepeg.org/enter_PCORE.html">http://gepeg.org/enter_PCORE.html</a> 
Provider	Research Group for Palaeoecological and Geoarchaeological Studies (GEPEG) at the University of Barcelona
Domain/thematic focus	Phytolith research
Geographic focus	Global
Type of data offered	Images of phytoliths recovered from different plant taxa and different geographical locations
Further information	<p>The PhytCore database is maintained by GEPEG, the Research Group for Palaeoecological and Geoarchaeological Studies at the Department of Prehistory, Ancient History and Archaeology, Faculty of Geography and History, University of Barcelona. The group's mission is to determine the physical and chemical composition of different archaeological materials (e.g., plant remains, bone, teeth and shells) and to analyse their age for the purpose of understanding site formation processes.</p> <p>The PhytCore database contains collections from GEPEG and other research groups of the University of Barcelona (including the Research Group for Palaeoecological and Geoarchaeological Studies), the Burke Museum of Natural History &amp; Culture in Washington (from the Department of Biology), and UW Biology.</p> <p>The phytolith images are collected from different sources and coded accordingly: modern reference plant material from the study areas (RC); modern soils collected from the same areas as modern plants or from areas that were previously described in terms of vegetation (SS); paleo soils samples (PS), and archaeological and paleoanthropological material (AS).</p>
Questionnaire ID	17

## **Assessment of the portal – strengths, weaknesses, innovative features**

### **Overall structure of the portal**

The evaluator describes the portal as “well-designed” and “easily searchable”. It provides the most important core data, and gives the possibility to upload one’s own record after registration.

Phytolith morphotype identification is highly “image-dependent”. The portal provides the functionality that multiple images can be uploaded for one single morphotype. This helps the identification and understanding of morphological variability of a particular morphotype. This procedure is a good practice that could also be envisaged in other use contexts.

### **Search functions**

The portal enables multi-layered search opportunities. Not only plant taxa or morphotype, but geographical location and archaeological sites can be searched throughout the site.

### **Mechanisms for downloading data**


Images can be viewed online, other data can be exported.

## **Conclusions for ARIADNE**

The following suggestions could be taken into consideration in the design of the ARIADNE portal:

- Good practice: Multiple images can be uploaded for a single morphotype. This helps the identification and the understanding of morphological variability of a particular morphotype.

#### 4.4.7 Scratchpads - biodiversity online

Website	<a href="http://scratchpads.eu">http://scratchpads.eu</a> 
Provider	Natural History Museum London
Domain/thematic focus	Biology (virtual research environment)
Geographic focus	Global
Type of data offered	Biodiversity of plant and animal species, specific research projects, in particular taxonomic information
Further information	<p>Scratchpads are an online virtual research environment for biodiversity, allowing anyone to share their data and create their own research networks. Thus, the service supports collaborative research. This can be preparing a paper with colleagues, building a bibliographic database or creating a reference collection of images and observations. Users who want to create their own scratchpad can sign up for free, set up and maintain their site, upload and annotate media files, and link them to taxonomic terms. Sites can focus on specific taxonomic groups, or the biodiversity of a biogeographic region, or any other aspect of natural history. Scratchpads are also suitable for societies or for managing and presenting projects.</p> <p>Key features of Scratchpads include: tools to manage biological classifications, bibliography management, media (images, video and audio), rich taxon pages (with structured descriptions, specimen records, and distribution data), and character matrices.</p> <p>Currently, there are about 620 Scratchpads by about 6,500 active users covering roughly 160,000 taxa in 975,000 pages.</p>
Questionnaire ID	21

## **Assessment of the portal – strengths, weaknesses, innovative features**

The evaluator took an exploratory approach, triggered by the wide variety of topics provided by the sites on this Scratchpads platform.

### **Overall structure of the portal**

The structure was assessed as “good”. There is a clear division into separate research projects (as a result of individuals or group efforts). Within each project, often the same sequence of possible topics is available (although personalization may be possible).

The evaluator recommends as good practice to have individual project archives, which are maintained by the respective contributor(s). Researchers thus remain responsible for their sites, with a personalized lay-out. Scratchpads provide an advanced web platform for digital collaboration beyond the core function of data archiving.

### **Search functions**

The general search box was considered very basic, but the faceted navigation and filtering of, for instance, literature is seen as very up to date and an example of good practice. The same holds true for the hierarchical taxonomy presentation with several tabs for each species (f.i. description, map, media, and literature). This creates a very clear and intuitive user interface.

### **Mechanisms for downloading data**

There is only limited functionality for downloading data, with proprietary file formats for individual data files with additional information (file share).

### **Options for uploading data**

Uploading data files is simply part of the website content management, without any specific functionality (file sharing). Advanced facilities for creating/sharing/linking taxonomy data (taxonomy editor) are available, with data import from a Microsoft Excel spreadsheet in the fixed ITIS standard or from the EOL web service. The documentation describes an export facility to a standardized NEXUS format (taxon-by-character data matrices/trees) and Darwin Core Archives (DwC-A) metadata standard (this was not tested, however).

### **Specific support functions & features offered**

A wide range of functions are made available. In a number of sites visited during this evaluation, these functions were present, but apparently used in a relatively limited way.

The help function was found to be excellent.


## **Conclusions for ARIADNE**

The following suggestions could be taken into consideration in the design of the ARIADNE portal:

- Good practice: the faceted navigation and filtering of literature
- Good practice: excellent help function



#### 4.4.8 SEDAC - Socioeconomic Data and Applications Center

Website	<a href="http://sedac.ciesin.columbia.edu">http://sedac.ciesin.columbia.edu</a> 
Provider	NASA EOSDIS and hosted by CIESIN, Columbia University
Domain/thematic focus	Socioeconomic data
Geographic focus	Global
Type of data offered	Spatial data with a focus upon socioeconomic factors such as health, climate, infrastructure and population. The portal also provides map tools to look at the data.
Further information	<p>SEDAC is one of the Distributed Active Archive Centers (DAACs) in the Earth Observing System Data and Information System (EOSDIS) of the U.S. National Aeronautics and Space Administration. SEDAC focuses on human interactions in the environment. Its mission is to develop and operate applications that support the integration of socioeconomic and Earth science data and to serve as an "Information Gateway" between the Earth and social sciences.</p> <p>SEDAC currently holds about 190 data sets. These can be searched by theme, year or format. Data and maps are available for download. There are also 35 data collections on quite heterogeneous thematic areas such as "Anthropogenic Biomes" or "Environmental Performance Index".</p> <p>A SEDAC User Working Group (UWG) provides ongoing advice and guidance regarding SEDAC activities and plans.</p>
Questionnaire ID	22

### **Assessment of the portal – strengths, weaknesses, innovative features**

A brief evaluation of the main functionalities and the usability of the site were carried out. The main conclusions from this test are:

- The portal has a clear and intuitive design and structure. There are clear headers which provide information about what they contain.
- There is a simple search function, but it could include more fields for search within. The portal lacks an advanced search function. It is not possible to search in a specific thematic area if one wants results which just contain data from Africa or Europe. Thus, the advantage is the simplicity, but it comes with a price – one may want (at least optionally) some further search fields and options.
- The download mechanisms are simple and direct. One just has to click on a link under the data downloaded. It was suggested that this function could also be available on the front page of each data set.
- The feedback and support for the portal is available on every page through its small web app, as well as a knowledge base. The feedback and FAQ are well structured and easy to understand and find in.

### **Conclusions for ARIADNE**


The analysis of this portal has unfortunately not delivered any specific good practices or recommendations that could be useful for ARIADNE.

## 4.5 Portals for research communication

### 4.5.1 Academia.edu

Website	<a href="http://www.academia.edu">http://www.academia.edu</a>
Provider	Academia.edu, USA
Domain/thematic focus	All topics
Geographic focus	Global
Type of data offered	Research papers
Further information	<p>Academia.edu is a platform for academics for presenting their research profile and sharing research papers. The company's mission is to accelerate the world's research.</p> <p>Researchers can use Academia.edu to share their research, monitor deep analytics around the impact of their research, and track the research of academics they follow.</p>
Questionnaire ID	26

### 4.5.2 Mendeley

Website	<a href="http://www.mendeley.com">http://www.mendeley.com</a> 
Provider	Mendeley Ltd.
Domain/thematic focus	n.a.
Geographic focus	Global
Type of data offered	n.a.
Further information	<p>Mendeley is an online service and database that supports different research processes and purposes. Users can register and set up a personal web page with their research profile on the Mendeley portal. The portal offers tools for: conducting initial research (search and discovery of papers, reading and analysis), writing a paper, review or grant proposal, submitting a dissertation for review, collaborating in project teams or lab groups, identifying collaboration partners, managing a curriculum, creating awareness, promoting oneself and publishing research results.</p> <p>The features offered include the “Reference Manager” (for generating citations and bibliographies), the “Read and Annotate” and “Add and Organize” features (for working with PDFs), the “Collaborate” feature and a feature for “Backup, Sync and Mobile”, and the “Network and Discover” feature, which offers access to millions of papers, one of the world’s largest crowd-sourced research catalogues. Papers can be searched as a free-text search or by disciplines.</p>
Questionnaire ID	26

### 4.5.3 ResearchGate

Website	<a href="http://www.researchgate.net">http://www.researchgate.net</a>
Provider	ResearchGate
Domain/thematic focus	All disciplines
Geographic focus	Global
Type of data offered	Papers, profiles of researchers
Further information	<p>The portal was founded in 2008 by physicians Dr. Ijad Madisch and Dr. Sören Hofmayer, and computer scientist Horst Fickenscher. ResearchGate today has more than 6 million members.</p> <p>ResearchGate offers users a personal web page with information about their institutional affiliation, skills and expertise, topics of interest, publication references or full-text (incl. views and downloads), co-authors on ResearchGate, followers and others they follow. ResearchGate automatically creates pages of institutions and departments with a list of publications, members, etc. This is based on the entities named in the researchers' profiles. The portal does not provide group functionality.</p>
Questionnaire ID	26

## Collective assessment of the three portals

### Users

The three research portals have a significant number of users who are active or otherwise interested in fields of archaeological research. The latter may be interested in closely related research fields, have a cross-domain perspective or a general interest in archaeology or specific sub-fields. The biggest attractor is Academia.edu with about 283,000 users. In comparison the 16,000 users on ResearchGate is a small but still substantial figure, while Mendeley attracts only small groups of archaeological users.

As of November 2014, more than 16,300,000 academics had signed up to **Academia.edu**, adding more than 4,500,000 papers and 1,300,000 research interests. Academia.edu attracts over 15.7 million unique visitors a month. Academia.edu at present has 282,824 users who take an interest in Archaeology or specific fields like Funerary Archaeology (11,607) or Bronze Age Europe (6505). In total 39,805 documents are referenced and a lot can also be downloaded (the exact figure of downloadable documents could not be identified).

**Mendeley** has about 29,000 literature references related to archaeology which have been uploaded by users. On the portal there are 83 "groups" interested in specific archaeological topics. Groups can be fully open, invite-only or not disclosed. 75 of the groups are open or invite-only, which can be followed also by non-members. In total there are 541 group members: 10 groups have 1 to 11 members, and 6 groups 11 to 21 members. Furthermore there is one group with 56 members and the largest has 66 members. Group members can add literature references and downloadable papers, post questions and comments, and receive alerts about such activities. For example the FAIMS - Information Management in Archaeology group has 6 members with 139 references/papers; the Roman Archaeology group 16 members with 544 references/papers.

On **ResearchGate** at present there are 15,894 users who follow the topic "Archaeology" and 12,595 are interested in one or more of 35 specific topics (not all also following the generic subject). Some figures of followers of specific topics are: 2825 Prehistoric, 1908 Environmental, 1032 Forensic, 589 Roman or 151 Aerial Archaeology. In total there about 10,000 archaeology-related literature references, some documents can be downloaded, but the exact figure is unknown.

One special feature of ResearchGate is the discussion function. Questions can be posted under one or more relevant topics of research. On 8 December 2014 there were 186 questions posed of which some triggered extensive discussion. For example the question "*What is the future of taking care of our Cultural Heritage?*" yielded a discussion thread of 77 responses.

Researchers who ask very specific questions receive answers by experts in the field. For example, "*A question for Marine Biologists - Can you recognize these bones*" (with a photograph of 20 small bones attached) received suggestions by five experts (e.g. "*it might be the maxilla of either fresh water fish or sea water fish*" or "*may be a humeri of sea turtles*")<sup>7</sup>.

### User profiles

All three portals offer researchers a personal web page to present their research profile, including institutional affiliation, research field/s, expertise, list of shared material, followers and others they follow, etc. Mendeley also allows setting up web pages for research groups. – Personal webpages are one of the most important features of the portals which could also be very relevant to users of the ARIADNE portal.

---

<sup>7</sup> Cf. [https://www.researchgate.net/post/A\\_question\\_for\\_Marine\\_Biologists\\_Can\\_you\\_recognize\\_these\\_bones](https://www.researchgate.net/post/A_question_for_Marine_Biologists_Can_you_recognize_these_bones)

## Information sharing

The portals mobilize large numbers of archaeological literature references, Academia.edu 40,000, Mendeley 29,000 and ResearchGate 10,000. Mendeley offers standard bibliographic data; on ResearchGate a large part of the references lack some bibliographic information, while on Academia.edu such information is largely missing. On Academia.edu a lot of material can be downloaded, while on the other two portals the largest part of the referenced publications cannot be accessed. The category data/dataset is missing on all three portals, hence data cannot be searched separately and it is also unlikely that a considerable number of references to data sources is available.

Suggestion: A useful feature for the ARIADNE portal could be to allow users to create a “MyData” or “OurData”-page which aggregates standard metadata for the data shared by the researcher or research group. Various related information could be ingested and presented using Linked Data.

## Groups

On Academia.edu, web pages for research groups can be created which aggregate information from their members' pages. Mendeley offers functionality for group management: “team plans” for a research institute or large project come with a considerable price tag, while some useful basic features for small open groups seem to be free of charge. Group members can add literature references, post questions and comments, and receive alerts about such activities. ResearchGate does not offer group functionality.

Suggestion: While functionalities for managing research groups may go beyond the primary purposes of ARIADNE, the “MyData” and “OurData” options with Linked Data would allow for populating dedicated pages with rich and interlinked information.

## Discussion function

Academia.edu does not offer a discussion function. On Mendeley members of groups can comment on posts of other members, while on ResearchGate questions can be posted under one or more relevant topics of research. Most questions receive comments, some expand to lively discussions. – Particularly interesting for ARIADNE are questions to experts in specific research fields concerning methods, objects and data sources. For example, researchers often seek help in identifying certain objects and post high-quality photographs for this purpose. One example where many such requests and expert answers can be found is the JISCM@il Archaeobotany forum.

Suggestion: ARIADNE could enable structured expert exchanges about objects in need of identification and thereby create a valuable database of objects and knowledge. The same might be possible for methods and data sources.



## 5 Lead users' suggestions for research portals

### ***An assessment of the current situation***

*"I haven't seen a portal which is exceptionally good. All have problems of structuring data after filtering/drop down menu choices and they do not allow a very refined filtering/drop down. So after an initial search you have to go through dozens of sites checking through the data, and after checking a lot you realise that they don't even have physical data to be downloaded. So it was a waste of time kind of thing. It should be made clear in the database where there are not physical data."*

### ***An ambitious portal vision***

*"A very ambitious vision: having a portal that links all (!) archaeological sites (where data are accessible), where one can search by location, timespan and so on and the data are linked with each other. You get (similar to Open Context) contexts with their associated finds, images linked with the contexts or the finds, information about scientific analysis (linked with the context resp. finds). You can download all data from a site (from one link) or if you only need the data of e.g. one context, then you can download these data only."*

*(Quotes from lead users' feed-back in the questionnaires)*

### 5.1 "Ideas and suggestions" (Module C) – concept and overall results

This section briefly describes the concept of Module C) "Ideas and suggestions" from the survey template (cf. Annex I), the background of the lead users, the portals surveyed, and the suggestions received from the participants.

#### **Concept of Module C**

In Module C) of the portal survey template the lead users were asked to suggest *"important and innovative features and functions which you would like to see in online archives and portals for researchers"*.

The module included four questions asking for:

- 3-5 features (e.g. specific services or tools) the lead user expects from an online portal which are or would be most helpful for their research activities. This could include features which are missing in current generation portals. In a table the lead user could describe briefly the suggested features and why they are/would be important to have.
- The lead user's general experience with search and other services on portals, including good solutions, main current bottlenecks, and possible improvements.
- Examples of useful services of portals the lead user could recommend, not including the portal evaluated in the survey.
- Any other suggestions, ideas and recommendations with regard to research data portals.

The lead users were invited to *"include both suggestions for simple improvements in details, as well as 'crazy' (visionary) ideas which are unlikely to be realised in the near future"*.

## Lead users' background and portals surveyed

Factors which will have influenced the suggestions of the lead users, especially the archaeological researchers among the survey participants, include their professional background, research focus and expertise as well as the choice of portal being surveyed.

### Lead users' professional background, research focus and expertise

- *Professional background:* The lead users were mainly archaeologists, 18 of 23 participants; the others five participants had a background in other disciplines (3) or were data managers (2). Of the archaeologists, 80% had at least 10 years of professional experience.
- *Research focus:* The research focus of the 18 archaeologists was quite diverse – both in terms of their geographic focus and research domains/specialties. The latter included e.g. Prehistory, Classical, Early Medieval, Landscape, and Urban, and specialties like Stone Age settlements, ancient agricultural technologies, funerary archaeology, and analysis of ceramics.
- *Experience in the use of digital archives/portals, tools and data:* 70% of archaeologists had used digital archives/portals for at least 5 years, 50% for 10 years or longer. Tools and data they created and/or used included remote-sensing tools, GIS, 3D and computational applications (e.g. statistical analysis), digital libraries and databases, geospatial data, material analysis data, metadata and conceptual knowledge; also various ICT applications for museums, archaeological sites and monuments were also mentioned.<sup>8</sup>

*In summary:* The strongest “bias” of the lead users group is that most were archaeologists. This was intentional because the ARIADNE data portal should mainly serve the archaeological research community.

### Portals surveyed

Also the sample of the portals surveyed (25) may have influenced the lead users' “wish list” of portal features. A portal was defined as a website that provides access to content/data of more than one organisation or project, including digital archives which curate third-party data. Most portals in the sample were “international”, i.e. provided access to content/data from research not only in one country. More specifically,

- 15 entities focused only or to a large extent on archaeological content/data: websites of digital archives (ADS, Arachne, DANS, MAPPA, OpenContext, tDAR), scientific databases (ceraDAT, CHARISMA, PhytCore), content/data federations (CLAROS, Fasti Online, Pleiades/Pelagios, research programmes (Mapping Death, Portal to the Past), and one community website (Bone Commons).
- 5 entities were state-of-the art portals of other domains: CIARD-RING (a registry of food & agriculture research information services and repositories), EUROSTAT, GBIF - Global Biodiversity Information Facility, Global Change Master Directory (earth & environmental data), and SEDAC - Socioeconomic Data and Applications Center.
- Two e-research environments: Morphbank (analysis and sharing of biological research images) and Scratchpads (biodiversity/natural history research, with a focus on taxonomy),
- Three academic/professional networking and content sharing platforms: Academia.edu, Mendeley and ResearchGate, which are used also by archaeologists.

---

<sup>8</sup> More information on the professional background, research focus and practices of the lead users is given in Chapter 3.

*In summary:* Most of the portals (15) had an archaeological focus and provided access to content/data from several projects. Some were websites of archives and research programmes with a pre-dominant focus on one country. Archaeological researchers are familiar with such websites and may expect that the ARIADNE portal provides similar as well as advanced online services (e.g. cross-search and exploration of several digital archives). The ten other entities were included to take account also of portals and e-research environments of other domains as well as what academic/professional networking and content sharing platforms offer to researchers.

### **Statements/suggestions**

Module C) yielded a total of 127 suggestions, mostly short statements but also extended description of the expected advantage of a suggested portal feature, good practice examples, and potential improvements of current generation portals.

The largest part (84) was suggested features (services, tools or other) of an online portal which would be most helpful for the lead users' research activities.

The further suggestions (43) comprised various examples of good practice as well as issues in current generation portals, potential improvements and other ideas and recommendations.

From this input we extracted and summarised the desired general and specific features of the ARIADNE data portal.

## **5.2 Suggested portal approach, services and specific features**

This section comprises 12 themes which were present in the lead users' suggestions, ideas and recommendations and important for the overall approach, design, services and specific features of the ARIADNE portal. These themes are addressed below by providing background, discussion and quotes of lead user statements.

### **5.2.1 Implement a good overview and navigation of resources**

This category corresponds to the overall wish of archaeologists to have an improved overview of existing data resources, i.e. data transparency. In the ARIADNE online survey, 95% of the respondents considered as very or rather important having a good online overview of available data.

In the context of the ARIADNE portal, the wish for data transparency applies to what may be found and accessed through the portal. This concerns the overall design of the portal which should be functional as well as attractive to use, e.g. with regard to overview, navigation, search options and other features.

Several lead users urged that a data portal should provide a very clear overview of what kind of data resources are available (e.g. sections for different data types), including statistical information on quantity and distribution (e.g. per provider, country/area, period, etc.). Directly related to the overview of data resources, the portal should make clear how the resources can be searched, accessed and (re-)used.

This form of information will also be appreciated within sections on specific data types. The information should be updated dynamically any time the metadata of new datasets are added to the underlying data.

**Recommendations:**

- Design a highly functional as well as attractive portal, e.g. with regard to overview of searchable content/data and portal navigation.
- Provide statistical information on the quantity and distribution of the data (e.g. per type of data, provider, country/area, period, etc.).
- Enable a good understanding of the data resources and how they can be searched, accessed and (re-)used.

**Selected statements:**

*"The combination of an appealing and – at the same time – functional design. The design of a portal etc. strongly helps the eye and mind decide what are the main features, which are the sections, where to go for help etc. Let's not think that what is scientific has to be boring but elegant!"*

*"Very few portals have a good design (e.g. how many steps are needed to reach requested information). There is no intuitive search possible, often there are very basic search formats (e.g. timeline search, map-based search, etc.)."*

*"There should be a map of the database and data structure with links so users can guide themselves in the forest of data."*

*"I would like to know without searching from which countries and organisations there are data in the portal".*

*"Generating quantifying and statistical information. – Quick assessment of the quantity of sites/artefacts/projects, etc. in a given area/collection/period, etc."*

*"For large portals (many datasets, data other than geo-referenced), a visual representation of statistics on portal content: is X concept well documented in the datasets available through this portal, and how well is X concept documented in the context of the datasets available in the portal. Quick gauge of whether the portal has relevant data and how much drill-down might be required to find an appropriate dataset."*

## **5.2.2 Ensure richness and added value of information**

Several of the lead user statements relate to the richness of the information envisioned to become accessible through the ARIADNE portal. These statements refer to the general documentation of sites (e.g. inventory, site assessment, reports, publications) as well as individual finds and their context (e.g. *"ceramic analysis documentation, zooarchaeology documentation, site documentation, description of graves, human anthropology etc."*).

Indeed, as shown by the results of the online survey, for the archaeological research communities all types of information/data are relevant (e.g. data from cultural heritage authorities, prospection and field survey, remote sensing, excavations, material and biological analysis, etc.)<sup>9</sup>. Within the ARIADNE data portal archaeologists will want to search and filter ("drill-down") to available rich data resources in various data formats as needed for their various research specialties and topics. The lead users' statements confirm this expectation (cf. the quotes below).

Interestingly, data quality was not emphasised that much. One statement addressed the topic directly: *"Detailed description of data provenance. - Possibility to evaluate data quality"*. Another

---

<sup>9</sup> Cf. ARIADNE First Report on Users' Needs. Deliverable 2.1, April 2014, pp. 79-81 (about 500 respondents); significant less demand was for data mining and model-based computing, which require large datasets that are not readily available in the sector.

lead user suggested: *"There should also be a sort of evaluation of the sites it [the ARIADNE portal] provides access to or at least a method to compare their basic features and data provided. The main bottlenecks are the lack of varied and rich documented content freely available."*

It is unlikely that the initial stock of data resources from ARIADNE partners will meet the high expectations from the ARIADNE portal. Therefore the project will have to seek integration of more resources from affiliated institutions as well as from other initiatives within and beyond the archaeology sector (e.g. remote sensing or molecular biology data). Possibly linking to data and publication resources not registered in the ARIADNE Registry may provide added value even if the own base of data is initially limited.

Also the European/international dimension of the ARIADNE resources may be an advantage with regard to attracting portal users. In the online survey 74% of the respondents considered "having access to international data(sets)" as very or rather important to conducting their research, while only 28% were very or rather satisfied with the current situation.<sup>10</sup>

Importantly, the users are generally aware of the current lack of openly accessible data in the archaeology sector, and will often be sufficiently happy to know where useful data exists and may be available upon request. In this regard, the ARIADNE portal should also actively promote a culture of open data sharing in the sector, through making users aware of the importance of sharing ("give and take") and pointing to guidance material and community repositories (including archives in the ARIADNE federation).

### **Recommendations:**

Potential users will expect an ARIADNE portal where they can search for various relevant data as needed for their research specialties and topics. It is unlikely that the initial stock of data resources from ARIADNE partners will meet the high expectations. Therefore the portal should:

- Integrate many resources from affiliated institutions as well as from other initiatives within and beyond the archaeology sector (e.g. remote sensing or molecular biology data).
- Emphasise the European/international dimension of the initial ARIADNE resources.
- Create added value through linking to data and publication resources not registered in the ARIADNE Registry.
- Actively promote a culture of open data sharing in the sector, through making users aware of the importance of sharing ("give and take") and point to good practice guides and community repositories (including archives in the ARIADNE federation).

### **Further selected statements:**

*"There are so many portals that each provide a small part of the information that I require for my research (high degree of fragmentation)."*

*"Well-structured help, frequently asked questions and clear guidance on using a portal is essential."*

*"Useful services include a calendar of events that could be of interest to the portal users and links to other relevant portals."*

*"Use of varied and rich multimedia material – Texts, images, plans, maps, videos, 3d representations are all features that can enhance the understanding of any topic. It is strongly advised to make available as much relevant material as possible in high quality."*

*"Multiple picture upload for each sample is vital. – Micro- and macro photos, object descriptions, drawings and photos are necessary for interpreting ceramic analyses data or any object data"*

---

<sup>10</sup> Ibid., pp. 110-111 (about 500 respondents).

*“Archaeological and scientific data should be in the same database not only for ceramics but for all material types.”*

*“Combining multiple archaeobotanical evidence in one single database site. – Would help the integration of the datasets and facilitate research interpretation.”*

*“It would be very important to combine databases with photographs in good quality. It would be also necessary to provide summaries/interpretations of certain data groups.”*

*“It should contain all the data types available on that object/site even when it is not accessible in the database and the contact person who has access to the data and may provide individual permission to access them should also be included. (...) If one finds a documentation type that exists but is not available online he/she can contact the data provider for individual access.”*

### 5.2.3 Help users understand and use specific terminology

The ARIADNE portal will serve all archaeological research communities as well as other users (e.g. cultural heritage administrators, citizens interested in archaeological topics). Lead users suggested that such a portal should avoid scientific language where possible and provide terminology aids if needed (cf. statements below). Indeed, it cannot be assumed that all portal users will understand the specific terminology of each archaeological research community/domain for objects and data production methods (e.g. remote sensing and surveying techniques, archaeometry methods, etc.).

The requirement of terminology aids concerns categories, concepts and terms used in search features but may also be taken account of in information pages. At a basic level, look up of glossaries or multi-lingual thesauri may be offered, or, more advanced, terms/concepts from such aids invoked dynamically or upon request. Ideally, multi-lingual terminology support is offered.

One lead user emphasised the need of minimum information requirements/standards for the description of scientific data; *“if one wants to upload scientific data there should be minimum requirements of that data; e.g. minimum requirements of ceramic thin section description etc. - This will allow better cross country/regional comparison of results”*.

Minimum information standards for scientific data, i.e. archaeometry and other analysis of various objects, seem to be widely missing, as confirmed by one ARIADNE partner that operates an archaeometry laboratory.<sup>11</sup> Terminologies are more standardised (e.g. International Code for Phytolith Nomenclature 1.0<sup>12</sup>) and followed in the research community, but may not be available in machine-readable formats or apply Linked Data principles. The standardisation of scientific information is of course a task of the scientific communities. However, providers of e-infrastructure like ARIADNE could promote the provision of terminologies according to standards<sup>13</sup> and in (semantic) web-based formats as required for machines<sup>14</sup> and humans (i.e. web-pages to look up term).

---

<sup>11</sup> In the biosciences many minimum information standards are available and included in the BioSharing platform, <http://www.biosharing.org/standards/> (filter on MIBBI Foundry).

<sup>12</sup> Madella M., Alexandre A. & Ball T. (2005): International Code for Phytolith Nomenclature 1.0. In: Annals of Botany 96(2):253-60, <http://aob.oxfordjournals.org/content/96/2/253.full.pdf>

<sup>13</sup> e.g. ISO 25964 - Thesauri and interoperability with other vocabularies, [http://en.wikipedia.org/wiki/ISO\\_25964](http://en.wikipedia.org/wiki/ISO_25964)

<sup>14</sup> World Wide Web Consortium (W3C) recommendations: Resource Description Framework (RDF), Simple Knowledge Organization System (SKOS), Web Ontology Language (OWL).

**Recommendations:**

- Provide aids for users not familiar with specific terminology/vocabularies (e.g. look up or dynamically invoked glossary or scope notes of thesauri).
- Consider terminology support within search features as well as information pages for data resources, where possible in multi-lingual form.
- Promote the provision of terminologies according to standards (e.g. ISO 25964) and in (semantic) web-based formats as required for machines and humans.

**Further selected statements:**

*“Use of simple and natural language – Avoidance to use scientific language where it is possible. There should be precaution for users that are not always familiar with specific terminology although they may belong to the sector. Moreover, the provision of a vocabulary might be of great help.”*

*“Multi-lingual comparative vocabularies for categories - Improving multilingual accessibility.”*

*“Terminology depiction – Especially when searching in databases in other countries it can be helpful to know what is meant by a certain term. Meaning can be very different, e.g. the term iron age comprises different data when used in Germany or in Scandinavia or GB.”*

*“Terminology assisted searches – The use of thesauri and other terminological technologies would improve the precision in data retrieval”*

**5.2.4 Integrate and link information resources**

The lead users thought of a data portal that demonstrates a high level of integration of and linking between information resources. The ideas referring to integration were informed by domain and cross-domain research databases, including the need for access to both data and publications, and the wish to consult websites of related projects and researchers (e.g. research profiles and credentials). Some selected examples were:

- Database-like integration: *“Archaeological and scientific data should be in the same database not only for ceramics but for all material types”.*
- Access to data and research papers, grey literature and other material: *“The content of a portal can be strongly supported and summarized by relevant scientific publications, reports, presentations, papers etc. freely available in the portal”.*
- Access to all related information: *“Integration through raw data, grey literature, laboratory datasets, open access literature and digital maps”.*
- Consultation of websites of related projects and researchers: *“Navigation - Direct link to the projects/researchers' websites”.*

The ARIADNE project will not store datasets and content (e.g. scientific images), but the portal will operate based on metadata of, and semantic relations between, data/content collections and items. Thereby the portal should provide virtual, database-like integration of resources.

Access to datasets and research papers, grey literature and other material can be considered as a very important feature of the portal. Therefore inclusion of metadata of document archives and publishers will be necessary. The increasing use of Digital Object Identifiers (DOIs) will help considerably to create and exploit links between publications and data.

One lead user envisioned searches across massive repositories of research reports and papers, confusing somewhat full-text and semantic search: *“Full text search on all words within all*



archaeological reports and articles, finding information on any subject (semantically)”; another suggested: “NLP - Text mining, to find relevant information inside documents.” This would require considerable pre-processing of material stored in digital archives, e.g. extraction, aggregation, indexing, etc.

Concerning archaeological reports specifically, often they are not published, although the reports may be the only record of the results of fieldworks and other investigations. Therefore enabling better access to such “grey literature” is one of ARIADNE’s objectives. The objective is addressed by exploring how Natural Language Processing (NLP) might be used to extract information from archaeological reports and semantically link the information with metadata of other content.

#### **Recommendations:**

- Make users aware that the portal does not store and create databases of primary data but operates based on metadata of, and semantic relations between, data/content collections and items.
- Provide integrated access to data and publications (i.e. include metadata from document archives and publishers).
- Specifically support the inclusion and linking of information (metadata) from archaeological grey literature, which may be produced with Natural Language Processing (NLP).

#### **Further selected statements:**

*“Links to other similar databases. - Access to different databases”*

*“Combining multiple ceramic or any other material analyses data in one single database site. - It would help the integration of the datasets and facilitate research interpretation*

*“Making available the reports of excavations (grey literature), to increase the information about the sites and the artefacts.”*

*“It would help if more grey literature and bibliographies could be made available in pdf format.”*

### **5.2.5 Follow and promote Linked Data principles**

Lead users considered also the importance of Linked Data for integrating information within the portal and linking to external resources. The statements addressed the potential of the Linked Data approach as well as the current lack of awareness of the benefits of such data; also the need of high-quality Linked Data was mentioned.

One lead user wrote: *“Linked data. Relations between data collections”*. Another elaborated the vision of *“having a portal that links all(!) archaeological sites (where data are accessible), where one can search by location, timespan and so on and the data are linked with each other. You get (similar to Open Context) contexts with their associated finds, images linked with the contexts or the finds, information about scientific analysis (linked with the context resp. finds).”*<sup>15</sup>

One lead user complained about the current lack of high-quality Linked Open Data: *“Linked Open Data that would be also useful for information sharing and exchange are often missing or are of bad*

<sup>15</sup> Open Context (OC), mentioned by the lead user, employs Linked Data principles, e.g. stable URIs of entities and descriptive properties, linked to one another through network graph relationships. Furthermore OC links to URI-identified concepts published by others, e.g. the Pleiades Gazetteer (places/locations) and the Encyclopedia of Life (biological taxa). The data records with stable URIs enable referencing and annotation of entities within OC as well as external applications; OC does not yet provide a queryable RDF triple store; cf. <http://opencontext.org/about/technology>

quality”; “Open” in Linked Open Data (LOD) means that the data is shared under an open license (e.g. Creative Commons Attribution – CC-BY) or released into the Public Domain.

Another survey participant emphasised that more promotion of LOD is necessary to leverage the value of existing resources: *“There is an absolute need to promote open linked data within the archaeological community and the mentality of its members, because without these data the significance of such portals is very much limited.”*

As a lead user example without reference to Linked Data, but a clear case for potential benefit of LD-based interlinking: *“Checking online resources there seems to be no combination of archaeological databases (e.g. ceramic) and scientific (ceramic petrography, geochemistry) databases. It seems that archaeologists and scientists develop databases for their own needs but the two don't meet. For example, ceraDAT contains geochemical data but no pictures and categorisations of vessels which were examined. Thus, ceramic technological data (composition) are not linked with archaeological data (vessel type, form, decoration), therefore it is difficult to interpret the results.”*

#### **Recommendations:**

- Deploy Linked Open Data (LOD) to integrate information within the portal and to link to external resources which follow LOD principles (e.g. HTTP URIs and RDF<sup>16</sup>).
- Demonstrate advantages of Linked Data to encourage further uptake of LOD principles by archaeological institutions and projects.
- Provide an LOD triple-store so that also external application developers can exploit resources of the portal for added value services (e.g. interlinking of databases).

#### **Further selected statement:**

*“I think that we should not only provide linked data, but be open to straightforward links to other online resources, preferably in a structured fashion.”*

### **5.2.6 Provide effective data search and filter functionality**

The lead users suggested multiple options for searching and filtering relevant information and datasets on the ARIADNE portal. The suggested features include:

- cascading drop-down menus/filters, based on various categories (“Filter selection. No need to ‘learn’ portal structures/biases – it’s presented.”)
- keyword based search, including auto-completion, various search operators, etc.,
- relevance ranking of search results,
- terminology assisted search (e.g. term lists, thesauri, etc.),
- faceted search and browse functions (possibly following initial keyword-based results), to narrow down and inspect results,
- map-based search and timelines (chronologies),
- semantic search functionality (assumed to allow more complex queries),
- content-based search, i.e. based on similarity of images, features of 3D models, etc.
- data recommendations, e.g. “related” resources or based on user activity (“People who looked at this file, also liked...”).
- options to save specific search results and combinations of search filters (“my searches”)

<sup>16</sup> Wikipedia: Linked data, [http://en.wikipedia.org/wiki/Linked\\_data](http://en.wikipedia.org/wiki/Linked_data)

Particular relevance was assigned to search based on maps and timelines (cultural chronologies). Clearly these are candidates for search capability most users would appreciate, not least because of the visualization these methods allow (addressed in the next section). One especially noteworthy suggestion was an “intermediate page” which would first present basic information about sites or particular types of finds, and then allow users to select the most relevant entries, compare them, select items for detailed study, etc. (the suggestion is included below).

Concerning the adequate set of search options there will be tensions between envisioned various search options vs. keeping it simple. A general guideline could be how many steps are necessary to actually reach possibly relevant data. Lead users expressed words of caution: *“Keeping it simple is the key. If people feel that it is too cumbersome, or takes too long in comparison with using a regular search engine such as Google the portal will not be used”*; and: *“Avoid over-complication with search tools. Avoid users being given too many functions all at once”*.

Following this advice we recommend to the portal developers to investigate the most relevant search options directly with members of the user community (including from the lead user panel) and regularly seek feedback on implemented solutions by the wider community.

It is also worthwhile to note that a high degree of integration and linking within registered ARIADNE resources as well as to various relevant external resources will add greatly to the perceived search capacity of the portal.

#### **Recommendations:**

- Investigate the adequate type and implementation of search options directly with members of the user community (various options should be suggested and scrutinized).
- Seek regular feedback on implemented solutions by the wider community.

#### **Further selected statements:**

There were about 40 suggestions on search, filter, browse and other features, including general as well as specific aspects, examples (some with screenshots of existing solutions). Below a selection of illustrative statements:

*“Bottlenecks apart from the query technologies are also to be found in the user interfaces that very often are difficult to understand/use.”*

*“Advanced search capabilities across different data types”*

*“Extensive search options – Allow truncating, allow search operators (e.g. “or”, exclude terms, etc.). – To help finding what I am looking for and narrowing down results.”*

*“Terminology assisted searches – The use of thesauri and other terminological technologies would improve the precision in data retrieval”*

*“Facet browse function – Will enable results of a keyword search to be narrowed down”*

*“Provide search option/filtering/drop down on documentation types (...) – In this way users can access data/documentations more directly and experts can access data more easily.”*

*“Several filtering options should be on one page and not in different tabs. Also, in order to fulfil as many user needs as possible several filtering categories should be available, apart from the free text search option.”*

*“Various types of search – by timeline, map, clustering, etc. – Increase chances to get relevant data in a more intuitive and accurate way”*

*“Map functionality and map selection – Especially in searches on research in foreign countries it would be helpful to have a map. It can also be helpful to choose data entries from a certain geographic area, especially when combined with chronological selection”*

*“Sort search results by relevance – Less precision and/or browsing through results needed”*

*“Semantic features – Interfaces for semantic queries would enable complex queries of the archives to get more rich and meaningful results”*

*“Similarity searches – Possibility to find similar archives/data/information based on query results in order to extend the research”*

*“Suggests resources you may be interested in based on resources you are looking at. – This will aid research by identifying similar resources you may not have considered”*

*“Useful features / best practices include for me: faceted search, integrated search, and being able to control/manage one's own search question.”*

*“Mobile apps - Facilitate and speed up the search.”*

*“A portal may contain links to diverse sites. An intermediate page, when selecting a link from the portal, would be useful if it provided some basic info about the site selected (type of material provided, culture, area or era of interest, data available for download or not etc.). A user could see some basic fields in the intermediate page or could skip it and go directly to the site. Moreover, there could be provision for selecting more than one sites and comparing them in this intermediate page.”*

### **5.2.7 Visualize data resources – Maps, timelines, and more**

Visualization has been addressed by several lead users and it seems obvious that this should be among the key features of the ARIADNE data portal. The visualization option mentioned most often in the context of searching and filtering of data are maps, followed by timelines (chronologies).

Map-based visualization depends on available geo-location data for sites and finds. Most archaeologists are familiar with location data and many projects use a geographic information system (GIS) to locate sites and finds and add available information (e.g. field observations, laboratory analyses of finds) to the database underlying the Web-GIS system. For older excavations only place names may be available (i.e. no exact coordinates). Chronologies depend on the dating of sites (and stratigraphic layers) based on the analysis of various site features and finds.

Map-based visualization was often suggested quickly to see if relevant data is available for countries, regions and further down to areas or certain locations (e.g., *“Quickly identify data for your study area”* or *“Quick gauge of how much data on specific locations”*). However, both location and date(-range) information will be required to establish the potential relevancy of available data for the study purposes.<sup>17</sup> Especially this is needed when searching across multiple countries, which is a core scenario for the ARIADNE portal, i.e. where the archaeologist will know well his/her study area but not sites/finds in other countries.

A major concern here is the importance of cultural periods/chronology. As one lead user suggested: *“Search should be available in absolute and relative date format because Neolithic in the Carpathian Basin and in Scandinavia was at a different time so researchers can avoid getting irrelevant targets when searching for periods.”* A date range based search, if too narrow, will not show available data

---

<sup>17</sup> Examples which provide this were the ARENA2 (<http://archaeologydataservice.ac.uk/Arena2/>) and Transatlantic Archaeology Gateway (<http://archaeologydataservice.ac.uk/TAG/www.jsf>) portals and the Erfgoed Breda website (<http://erfgoed.breda.nl>).

for a culturally defined period (e.g. "Iron Age" or "Roman") in different countries/regions. Therefore also named period based ("*relative date*") search should be available.

It may be worthwhile to note that maps and chronologies do not allow visualizing relations between sites/finds beyond distance in space and time (period). However, there was not much consideration of other visualizations except that one lead user thought of tools for "*exploratory data analysis utilising good graphical visualisation*", and that there was frequent mention of links/linking and Linked Data. Therefore here we suggest considering also Linked Data (LD) based visualizations, like the density and web of links of the LD graph around concepts.

In general, visualization on maps, with timelines, based on linkage will allow portal users to dig deeper into clusters of data resources as shown when entering the ARIADNE portal.

#### **Recommendations:**

- Offer search & filter functionality based on maps as well as date-ranges (timelines) & named periods. Both features will be required to allow users to select potentially relevant datasets.
- Consider also visualizations based on Linked Data (e.g. density and web of links of the Linked Data graph around concepts).

#### **Further selected statements:**

*"A spatial portal could be interesting for several research areas since it could give consistency over disciplines as well get more detailed data over time."*

*"Map Interface for searching with search extent polygon tools – Quickly identify data for your study area which you will know".*

*"Filtered/searched results should be presented on a map. - There should be a 'gradual map' (...) visualised in order to understand the results better or provide further ideas for refining search results".*

*"Map functionality and map selection – Especially in searches on research in foreign countries it would be helpful to have a map. It can also be helpful to choose data entries from a certain geographic area, especially when combined with chronological selection".*

*"Enable visualization based geo-spatial accuracy of the data – whether archaeological sites or artefacts have precise or approximate coordinates".*

*"Advanced timeline search. Possibility to select data based on temporal criteria would allow retrieval of information concerning specific period".*

*"Multilingual and multi-local timescales. For Fasti, we have timescales for each country which return, reciprocally, start and end dates for each period ('iron age', Mauretanean'). Even on a countrywide basis this is difficult, as 'Classical Greek' does not even apply to the whole of Sicily, much less Italy. In Spain it is proving a massive problem, as the coast and the interior have very different trajectories, but it is a subject to address."*

## 5.2.8 Provide data preview and license information

### Data preview

When a portal user has discovered some interesting data resources, e.g. a list of narrowed down search results, they will want to check if they are indeed relevant for the study purpose. To allow this check, lead users suggested offering a data preview feature, e.g. *"It would be of help if there could be a preview popping up of the data to be downloaded, to see if it is useful to own research interests"*. Mechanisms which allow a fast way of assessing the relevancy of resources (e.g. snapshots, "look inside" functionality) would certainly be very welcome.

### License information and re-use of data

In data previews or before downloading data it will also be important for portal users to see the license under which the provider makes the data available. This is important especially if they want to re-use the data, e.g. include data in a dataset or use images on a website (for example, to compare specimens). Re-use of data for new research is a major argument in "open data" mandates of funding agencies as well as for the need of e-infrastructure. Little is known, however, about data re-use in archaeological research.<sup>18</sup>

Lead users rarely addressed IPR and licensing issues, although there were statements implying re-use of data, e.g. *"Source Data integration. Tool that would allow for extraction of raw data from several (dispersed) project archives or thematic portals into a new table"*.

One lead user, who works on a large governmental project that brings together data from many sources, mentioned, *"formal agreement with the provider; IPR issues, selection of the CC license"*; licensing is part of the regular workflow of this lead user. Others addressed licensing under a survey question on problems in the use of online archives and databases (e.g. *"unclear licensing models or missing licensing"*) as well as in the evaluation of a portal.

While in this study IPR/licensing received little attention, in the online survey it was a burning issue (cf. the selected statements in Section 3.2.4). We understand making portal users aware of available (or missing) licensing information as an important aspect of the user-friendliness of data services. Furthermore, it can be an instrument to promote "open data" policies. Ideally the portal would allow users to filter available data according to restrictions on allowed usage.

### Recommendations:

- Implement a data preview mechanism that enables portal users to check if discovered data resources are actually relevant for the study purpose.
- Make users aware of available (or missing) license information of the data provider, e.g. to identify restrictions which impede re-use.
- Enable portal users to filter available data according to allowed usage, from Public Domain to fully (c)-restricted.

### Further selected statements:

*"Access (Visualization)/Download of all data "Possibility to look at the data beforehand, download of data to link them with own data."*

*"Preview data (e.g. thumbnails). To scan the data, to see what is relevant in a quick way."*

---

<sup>18</sup> One available study is Faniel, Ixchel *et al.* (2013): The Challenges of Digging Data: A Study of Context in Archaeological Data Reuse. JCDL 2013 proceedings (preprint), <http://www.oclc.org/content/dam/research/publications/library/2013/faniel-archae-data.pdf>

*“Unclear licensing models or missing licensing lead to problems when it comes to publication”*

*“Within (...) we are developing this kind of services on behalf of different users (...), always in full compliance with licensing, copyright/copyleft and use/re-use official policies of our Ministry and other relevant laws”.*

*“Good to have guidance for all potential contributors on the areas of metadata creation, licensing, standards but they are not really clearly laid out within this portal”* (a search portal for earth and environmental data).

*“The portal is organized quite well. Thanks to the various sections is easier to navigate in a direct way to find the information you are searching for. (...) Data is available under the CC-BY-NC-ND 3.0 license”* (a portal presenting field survey and excavation data; note: the license does not allow any re-use of data, “ND” [NonDerivative]).

### **5.2.9 Support different download options and external applications**

This section summarises lead users' suggestions concerning data download as well as services a portal should provide to external applications. The upload of data has been mentioned seldom but is also addressed in brief.

#### **Upload of data**

Surprisingly, in Module C) “Ideas and Suggestions” the upload of (meta-)data was addressed seldom, although this was one item in the portal evaluation. The ARIADNE portal will not invite data deposits, but its success depends in good part on richly filled underlying repositories. Therefore the portal should promote a culture of open data sharing based on “give and take”.

One lead user envisioned the portal or underlying repositories filled with a lot of visual material: *“Multiple picture upload for each sample is vital. Micro- and macro photos, object descriptions, drawings and photos are necessary for interpreting ceramic analyses data or any object data.”*

Another lead user expected a data portal to support *“image search using a keyword or uploading a file”*. The latter implies content-based search based on automatic image comparison (which the ARIADNE portal may or may not offer), but it would not require the portal to store the uploaded image.

#### **Download options**

Several lead users emphasised that a portal should offer different download options. This included single item and bulk download as well as export/download in different open formats, for example:

- *“Not only download for a single data set, but for a whole bunch. Allows working offline.”*
- *“Download of the information in modifiable format. Easy input and storage of the downloaded data for personal databases.”*
- *“Making available to download CSV and XML files.”*

In this context also tools which support data conversion and merging were considered, but not necessarily to be provided by the portal.



## Services for external applications

Lead users also considered that a portal should enable external applications to exploit available data, metadata and conceptual knowledge resources. This could include a well-documented Application Programming Interface (API), OAI-PMH target<sup>19</sup>, SPARQL endpoint<sup>20</sup> or other means.

The lead users mostly thought of an API or other “tool” and most mentions referred to geographic information, especially location data. Some selected statements are:

- *“API. So we can integrate with our internal tools and services”*
- *“Provide tool or an API to existent tools to easily represent geographic information on maps for selected data.”*
- *“For GIS applications wms [Web Map Service] is of big help because with it you can integrate gis data from foreign sources into your own analysis without downloading.”*

One lead user suggested RSS (Really Simple Syndication) feeds (*“Alert /Subscription system – Inform me when new data sets in my interest are available (RSS & Atom)”*). Such feeds could indeed be very useful to provide services that enrich the information of websites of research communities in particular subjects or geographic regions (e.g. alerts on new datasets in digital archives).

Lead users also suggested that the portal could use external services, for example, services of geospatial data infrastructure or novel Cloud based services and tools.<sup>21</sup>

### Recommendations:

- Support download of data/metadata, including single item and bulk download as well as export/download in different open formats (for not directly accessible data refer to the download page of the data repository).
- Provide interfaces to allow external applications to exploit available data, metadata and conceptual knowledge resources (e.g. a well-documented API, OAI-PMH target, SPARQL endpoint).
- Consider also RSS feeds for researchers and websites of research communities in particular subjects or geographic regions (e.g. alerts on new datasets in digital archives).

### Further selected statements:

*“Lack of clarity on how to obtain or download data from a portal or web service is frustrating.”*

*“You can download all data from a site (from one link) or if you only need the data of e.g. one context, then you can download these data only”.*

*“To allow users to download a complete collection of data regarding each archaeological context/site/object”*

*“Tools to view and/or convert data e.g. map from one metadata schema to another. Or merge multiple xml files into a single table. This is very important for systems using their very own conventions.”*

*“Geolocation – This is a fundamental activity in order to allow users to search data through spatial position and also to possibly overlap web portal data and anyone's own data.”*

---

<sup>19</sup> Open Archive Initiative – Protocol for Metadata Harvesting, <http://www.openarchives.org/pmh/>

<sup>20</sup> SPARQL Query Language for RDF, <http://www.w3.org/TR/rdf-sparql-query/>

<sup>21</sup> e.g. GIS Cloud, <http://www.giscloud.com>

*“Advanced integration with geospatial data infrastructures/repository/catalogues (as for Italy ‘Portale Cartografico Nazionale’ or local public online map bases) – To search fundamental and useful map bases necessary to better geo-reference and represent geospatial archaeological data.”*

### **5.2.10 Provide personalized information services**

Lead users suggested that a portal should provide personalized services which ranged from relatively simple mechanisms to quite demanding processes. Specifically mentioned were:

- a “save searches” feature, which also provides sharable links to searches,
- notifications/alerts, e.g. e-mail list or RSS feed subscription,
- filter relevant information based on tracking users’ interests,
- a dashboard which allows users to record use of data and services/tools,
- “MyData”/“OurData” pages for researchers and research groups, and
- support (or provision) of content management (e.g. Wordpress, Drupal).

This adds service features to the already long list of suggested search and filter options as well as new services (e.g. support or provision of content management systems). A “save searches” feature and alerts are likely to be good candidates for service provision. Tracking users’ behaviour on the portal and offering pre-filtered information may not be appreciated by many users.

“MyData”/“OurData” pages were envisioned to aggregate metadata for data which a researcher or research group have shared through open access repositories; furthermore Linked Data would allow for populating these dedicated pages with other interlinked information, selected by the researchers (e.g. research profiles on institutional websites or professional networking portals).

As recommended with regard to the multitude of suggested search and filter options, the portal developers should investigate, select and implement relevant services in collaboration with members of the user community.

#### **Recommendations:**

- Implement personalized information services which meet evaluated clear needs of large user segments.
- Prioritise features which support users’ control of specific searches, results and use of data.
- Consider notifications (alerts), e.g. e-mail lists or RSS feeds, as good candidates for personalized services.

#### **Selected statements:**

*“Being able to control/manage one’s own search question”.*

*“Possibility to save specific searches”.*

*“Save Searches option – Will help manage data and resources found using the portal”.*

*“Sharable persistent links to searches – Collaboration and communication”*

*“Alert/Subscription system. Inform me when new data sets in my interest are available (RSS & Atom)”*

*“Notifications/radar. Staying informed on new entries to the database”*

*“Possibly include a dashboard for users which provides summary of tools and data used”*

*“CMS integration – Integrate using Wordpress, Joomla, Drupal”.*

*“Through regular use the portal will understand what areas of the world my research is usually located and what data themes I am interested in (...). Record users' trails through data”.*

### **5.2.11 Enable linking and exchange of professional information**

Lead users suggested that the portal could enable expert discussion and answers to research questions (e.g. identification of finds). Furthermore linking to information of academic/professional networking platforms was considered, e.g. Academia.edu which is used by many archaeologists.

Academia.edu and other platforms (Mendeley, ResearchGate and others) offer researchers a personal web page to present their research profile, including institutional affiliation, research field/s, expertise, list of shared material, followers, and others they follow, etc. The personal web page within a large community portal is one of the most important features of these portals. Furthermore the sharing/promotion of research literature (references, downloadable material) and the opportunity to pose questions or contact other researchers are appreciated.

The ARIADNE portal would certainly not aim to establish a similar platform but could benefit from the available information. A clear candidate is the research profile of researchers. The portal could allow researchers to link their profile to own data records or, more effective, create a “MyData” page which aggregates records automatically and is linked to the research profile.

A clear need of many archaeologists is to ask colleagues for advice on matters like specific methods, objects or data sources. This is mostly done through direct contacts or within a larger circle of specialists. Listserv forums are frequently used to pose questions, but do not produce a structured knowledge base. The same applies to discussion threads on ResearchGate. Websites which invite community members to upload images of objects in need of identification (e.g. Paleobot.org<sup>22</sup>) often show little growth in the number of entries. To sum up: If it is considered that the ARIADNE portal supports expert discussion, the most effective approach and tool should be investigated thoroughly.

#### **Recommendations:**

- Enable linking of existing profiles of researchers on institutional websites and professional networking platforms to portal web pages (e.g. “MyData” pages).
- Investigate how the portal might enable effective expert discussion.

#### **Selected statements:**

*“A discussion blog/forum (on the home page). The possibility to find answers to my questions and tips for my work.”*

*“At the moment, Research Gate and Academia.edu are the most suitable online resources for getting scientific data and articles and it is easy to contact with researchers and ask for additional information/data if needed. A very useful feature of both is that you can ask questions that are posted to researchers who have similar interest to you and they provide answers and publications for your research.”*

*“Personal webpages are one of the most important features of professional networking portals which might also be relevant for the ARIADNE portal.”*

*“Navigation – Direct link to the projects/researchers' website”*

---

<sup>22</sup> Paleobot.org is intended to support the identification of archaeobotanical specimens, <http://www.paleobot.org>

*“ARIADNE could enable structured expert exchanges about objects in need of identification and thereby create a valuable database of objects and knowledge. The same might be possible for methods and data sources.”*

### 5.2.12 Support online research work (e-research)

Several of the lead users suggested tools for e-research, which means research performed primarily online. Visions of Virtual Research Environments (VREs) see researchers carrying out their work fully online, supported by powerful integrated tools, dedicated services and access to any kind of data and in whatever volume (e.g. “big data”) the researchers may need for the research. However, typically e-research does not span the whole workflow of researchers but only one or a few tasks, and the researchers are happy enough if they can carry them out more effectively with tools and data not available locally.

A portal which not only allows searching and downloading of data but also supports some other research tasks could be quite exciting for some archaeological researchers. The lead users envisioned e-research tools for several research tasks which include: mining, extraction, integration, measurement, comparisons, statistical and other analysis of numeric data or digital surrogates of research objects. Some tasks and tools were described generically while others elaborated in detail, for example:

- *“Online tools for analysis of digital content. – Remote access and analytical analysis”*
- *“Data warehouse functionalities. – Enable execution of common data analysis tasks.”*
- *“Source data integration. Tool that would allow for extraction of raw data from several (dispersed) project archives or thematic portals into a new table (for ex. provide me with the available isotope values (13C/15N) for bone from the Neolithic period).”*
- *“Comparison tools, statistical tools. – Scientific use of digital data”*
- *“Online measurement tools (e.g. digital calliper). – Improve interaction with digital content by being able to perform measurements on digital objects.”*
- *“Combining multiple ceramic or any other material analyses data in one single database site. – It would help the integration of the datasets and facilitate research interpretation”.*

The examples suggest two ways in which a research portal can support e-research tasks: tools which enable the combination and integration of data and, building on this, tools to process and analyse the data. With regard to the data, there can be existing numeric data (e.g. isotope values), or the values must first be produced with a measurement tool (which is also possible from images). Given the numeric data, statistical analysis or simply comparison of values can be done.

While research conducted with such tools is advanced e-research work, many researchers would certainly appreciate if they could access and work with various content/data related to their research questions in a more integrated way. For example, in urban archaeology they would appreciate *“integration of raw data, grey literature, laboratory datasets, open access literature and digital maps by means of a webGIS platform (...) and observing and analyzing them from and towards different point of interest and informational objects”.*

We assume that for enabling e-research the ARIADNE portal will in the first place have to support “integration” of content and data through integrated access to resources which relate in one way or other to the objects (archaeological sites, finds or other) and questions of archaeologists. Indeed, lead users stressed the situation that they have to search and collect from many websites parts of the different content/data they need for their research. For example, one lead user specifically noted

a lack of “integration” or “linking” of archaeological and archaeometrical research data and analysis (cf. the selected statements below).

**Recommendations:**

- Support integrated access as required for studying various research resources online (e.g. linking and comparing content).
- Provide or link to tools which enable researchers to extract and combine data (e.g. images from different databases, numeric data to produce a derived dataset).
- Provide or link to tools for data processing and analysis (e.g. statistical analysis, image data processing and analysis).

**Further selected statements:**

*“There are so many portals that each provide a small part of the information that I require for my research (high degree of fragmentation).”*

*“I have used several other social science portals: A problem is often that you have to go to several portals to find different parts of data. A spatial portal could be interesting for several research areas since it could give consistency over disciplines as well get more detailed data over time.”*

*“Combining multiple ceramic or any other material analyses data in one single database site. – It would help the integration of the datasets and facilitate research interpretation. For example, the ceramics has xy composition which is not characteristic of that area but some 100 kms away there are such deposits. Checking the stone tool assemblage of that site (in the same database) there are stone tools which also point to an area some 100 kms away so ceramic and stone trade from this area can be assumed.”*

*“Archaeological and scientific data should be in the same database not only for ceramics but for all material types. (...) Checking online resources there seems to be no combination of archaeological databases (e.g. ceramic) and scientific (ceramic petrography, geochemistry) databases. It seems that archaeologists and scientists develop databases for their own needs but the two don't meet. (...) Thus, ceramic technological data (composition) are not linked with archaeological data (vessel type, form, decoration), therefore it is difficult to interpret the results.”*

## 5.3 Good practice examples, experiences, opportunities for improvement

### 5.3.1 Good practice examples and general experiences with portals

After they had carried out the evaluation of a specific research portal, we asked the members of the lead user panel if they could recommend any good practices from digital archives and portals other than the one they had just looked at. The lead users suggested some portals as role models and highlighted some features which they found particularly useful in their work. The following quotes are a representative sample of the feed-back obtained.

- *"For me, role models for good data portals are 'open context' and tDAR; the only thing missing here is the wms."*
- *"The search option is a powerful tool and it should always have a simple and more advanced version with multiple criteria. There should also be a sort of evaluation of the sites it provides access to or at least a method to compare their basic features and data provided. The main bottlenecks are the lack of varied and rich documented content freely available."*
- *"Useful services include a calendar of events that could be of interest to the portal users and links to other relevant portals."*
- *"There is a portal for earth and environmental science, PANGAEA, which provides pretty useful advanced search tools (<http://www.pangaea.de/advanced/>). When you want to work with a large amount of data from a portal this can save a lot of time."*
- *"A good practice examples is <http://www.wolfram.com/data-science-platform/>*
  - *Wolfram Data Science Platform lets you use data sources that are structured or unstructured, and static or real-time. See at the web site for more details <http://www.wolframalpha.com/>*
  - *Wolfram|Alpha introduces a fundamentally new way to get knowledge and answers, not by searching the web, but by doing dynamic computations based on a vast collection of built-in data, algorithms, and methods. See at <http://www.wolframalpha.com/about.html> for more details."*
- *"A good practice example is: <http://www.giscloud.com/data-publishing> -for-web-and-mobile. GiSCloud allows users to search, edit, publish and share interactive maps; some map browsers and viewers allow users to search, visualize and explore media enriched data"*
- *"It is very important to develop new data delivering services in order to allow users to search also semantically and spatially, filter, collect, organize, download and re-use public archaeological data through public portal supplied with webGIS platform and Cultural Resources Management functions. Within SITAR Project we are developing this kind of services on behalf of different users, from specialists to digital creatives, up to citizens and tourists, in order to better deliver data and related digital objects, ever with a full compliance with licensing, copyright/copyleft and use/re-use official policies of our Ministry and other relevant laws."*
- *Good practices are the timeline function available in the TAG (<http://archaeologydataservice.ac.uk/TAG/www.jsf>) and ARENA2 (<http://archaeologydataservice.ac.uk/Arena2/>) portals.*
- *"The ability to rank search results by different criteria like in the DANS (<https://easy.dans.knaw.nl/ui/home>) portal."*





- *Useful features / best practices include for me: faceted search, Integrated search, and being able to control/manage one's own search question"*

### 5.3.2 Things to be improved

We then asked the members of the lead user panel about concerns, challenges and problems which they experience in their work with digital archives and portals, irrespective of the specific portal for which they were asked to provide an evaluation. The following statements are taken from this section. They document very well the general problems that still need to be addressed in the future.

- *"Search functionalities can be improved on most of the systems I used so far. Often there is only a simple search which then can be narrowed down by facets. This is not always enough. I wish the archives, databases and portals would make use of search functionalities and download and export options as can be seen on library catalogues."*
- *"Very few portals have a good design (e.g. how many steps are needed to reach requested information). There is no intuitive search possible, often there are very basic search formats (e.g. timeline search, map-based search, etc.). The purpose of portal should be clearly stated (e.g. administrative, research, communicative, etc.)."*
- *I haven't seen a portal which is exceptionally good. All have problems of structuring data after filtering/drop down menu choices and they do not allow a very refined filtering/drop down (the picture below provides a good example of very useful filtering options) so after an initial search you have to go through dozens of sites checking through the data and after checking a lot you realise that they don't even have physical data to be downloaded so it was a waste of time kind of thing. It should be made clear in the database where there are and there are not physical data.*
- *Several filtering options should be on one page and not in different tabs. Also, in order to fulfil as many user needs as possible several filtering categories should be available, apart from the free text search option. A not-too-good example in this respect:*  
[http://www.heritagegateway.org.uk/Gateway/Advanced\\_Search.aspx?reset=true](http://www.heritagegateway.org.uk/Gateway/Advanced_Search.aspx?reset=true)
- *Improvement should be in terms of completeness of data and various download formats.*
- *"Where drop-down lists are used well they are very helpful, but restrictions to select only one term from a list can be frustrating. Too many or too few required fields on advanced searches can hamper the user depending on context. A free text field alongside advanced/faceted search is flexible for the user. Lack of clarity on how to obtain or download data from a portal or web service is also frustrating."*
- *Existing portals in general only seldom are able to provide advanced features to simplify the retrieval of relevant results and very often they force the user to manually select and filter the material they provide. Bottlenecks apart from the query technologies are also to be found in the user interfaces that very often are difficult to understand/use.*
- *Linked Open Data that would be also useful for information sharing and exchange are often missing or are of bad quality.*
- *There are so many portals that each provide a small part of the information that I require for my research (high degree of fragmentation).*
- *Have used several other social science portals. A problem is often that you have to go to several portals to find different parts of data. A spatial portal could be interesting for several research areas since it could give consistency over disciplines as well get more detailed data over time.*

- *Typically, the main bottlenecks and weaknesses in archives and portals are: presenting the selected query is not always logical, sometimes you lose the original question.*
- *I would welcome accelerating and facilitating search on portals and easy tutorials of how to use the portals.*
- *"It would be helpful if there was more transparency about the data available on a portal before one starts a particular search, for example with diffusion maps of the sites/regions from which data are provided, and with information about the archaeological "epochs" that are covered."*

## 5.4 Other suggestions

Finally, we asked the lead users if they had any further suggestions on how research portals and the access to digital data could be improved in the future, and if they had any recommendation for portals (other than the one they had evaluated) which could serve as best practice for ARIADNE. They made the following suggestions:

- *"At the moment, Research Gate and Academia.edu are the most suitable online resources for getting scientific data and articles and it is easy to contact with researchers and ask for additional information/data if needed. A very useful feature of both is that you can ask questions that are posted to researchers who have similar interest to you and they provide answers and publications for your research. Another good feature of academia.edu is that you can open a session on your work and others with similar interest are invited to comment on your paper, revise specific parts and provide further data that you might have missed. This is very useful as researchers from all over the world can add to your work quickly and efficiently."*
- *"There is an absolute need to promote open linked data within the archaeological community and the mentality of its members, because without these data the significance of such portals is very much limited."*
- *"Avoid over complication with search tools. Avoid users being given too many functions all at once."*
- *"Well-structured help, frequently asked questions and clear guidance on using a portal is essential. A map search facility can also be very useful as seen in many portals."*
- *"Keeping it simple is the key. If people feel that it is too cumbersome, or takes too long in comparison with using a regular search engine such as Google the portal will not be used."*
- *"Scientific methods should be harmonized or give more options for data recording not with only one method. It would be very important to combine databases with photographs in good quality. It would be also necessary to provide summaries/interpretations of certain data groups."*
- *"A very ambitious vision: having a portal that links all(!) archaeological sites (where data are accessible), where one can search by location, timespan and so on and the data are linked with each other. You get (similar to OpenContext) contexts with their associated finds, images linked with the contexts or the finds, information about scientific analysis (linked with the context resp. finds). You can download all data from a site (from one link) or if you only need the data of e.g. one context, then you can download these data only."*
- *"The real issue, in the end, is content. I am not alone in thinking that Pleiades has a very meagre content, or that Oasis is enormously rich. But both allow you to get at the content in*

*a fairly straightforward fashion. I DO think that we should not only provide linked data, but be open to straightforward links to other online resources, preferably in a structured fashion.”*

- *“It would help if more grey literature and bibliographies could be made available in pdf format.”*
- *“The time period for the data within this portal is very short, but could be very useful for the whole research community if the time period could be extended to cover more of our history.”*
- *“A comment function to get in contact or to provide useful information (e.g. about errors or about additional data sources) would help to qualify the data and to reduce the burden of trial-and-error in searches.”*

## 6 Conclusions and recommendations

The ARIADNE First Report on Users' Needs (D2.1, April 2014) allows a good understanding of the needs of the archaeological research and data management community with regard to e-infrastructure and services. The results of the online survey provide an especially solid basis for the general approach of the ARIADNE data portal, which will be the main interface for the end-user.

The survey results confirmed that researchers in archaeology and related heritage sciences in Europe (and beyond) lack services that enable them to discover, access and (re-)use data needed for their research. Asked about which services would be very important or helpful for data portal, 80% of the about 500 researchers surveyed wanted a good overview of available data as well as capability to search data stored in different digital archives; 63% also expected innovative and more powerful search mechanisms.

However, the online survey was aimed to produce a broad overview of the current situation with regard to archaeological research data including, among other topics, data production, deposit and sharing/publication. Therefore the survey did not cover specific requirements which could guide the development of the ARIADNE data service portal.

The recommendations of this Second Report on Users' Needs are now meant to support the development of the overall approach, design, services and specific features of the ARIADNE data portal. The recommendations have been extracted and summarized from suggestions, good practice examples and ideas of 23 "lead users", researchers in archaeology, cultural heritage sciences and data management. They served as lead users based on their professional background as well as experience in the use of existing online archives/databases and other community portals. Before they gave their suggestions and ideas, each lead user had also reviewed at least one relevant portal in the field of archaeology or another discipline.

First, we provided a high-level view of the recommendations, using and adapted Kano model.<sup>23</sup> This view is meant to support the understanding and discussion of the recommendations, as well as decision-making on the development of the ARIADNE data portal. The next section presents the full list of recommendations. The final section provides ten conclusions concerning the development of the ARIADNE data portal.

### 6.1 High-level view of the recommendations

This report presents 12 sets of recommendations for the overall approach, design, services and specific features of the ARIADNE portal (in total 34 recommendations). Each set of recommendations corresponds to one section in Chapter 6.2, which provides background, discussion and quotes of lead user statements. The high-level view groups the 12 sets according to the following categories:

- **BASIC:** This category comprises service requirements which the portal must fulfil to be perceived as a valuable resource for the archaeological research community. The requirements include a good overview and navigation of data resources; richness and added value of information, and effective data search and filter functionality.
- **SUPPORT:** This category comprises services which are less critical but very helpful for portal users to understand, evaluate, download or use the data online (e.g. with/for external

---

<sup>23</sup> A technique for deciding which features should be available in a product or service, developed by and named after Dr. Noriaki Kano in the 1980s. In this technique, user requirements are grouped into different categories depending on how essential they are for user satisfaction, and in terms of how well the current solution satisfies this requirement.

applications). This includes support with regard to terminology, easy evaluation of resources (e.g. preview, license information), and actual data access.

- **ENRICH:** This category comprises services which can leverage the portal resources with regard to richness of information and integration/linking both within the portal and with external resources (e.g. through Linked Data). These services build upon and extend the BASIC service requirements.
- **EXCITE:** This category comprises services which can excite broad segments of portal users (e.g. visualization and personalized services) as well as users who expect advanced support in online research work (e-research). These services offer users enhanced or additional functionality.

The sets of recommendations which correspond to the four categories are shown in Figure 6-1.

**Figure 6-1: Recommendations for research portals (in summary of an evaluation of portals by lead users)**

<p style="text-align: center;"><b>ENRICH</b></p> <ul style="list-style-type: none"> <li>○ Integrate and link information resources</li> <li>○ Follow and promote Linked Data principles</li> <li>○ Enable linking and exchange of professional information</li> </ul>	<p style="text-align: center;"><b>EXCITE</b></p> <ul style="list-style-type: none"> <li>○ Visualize data resources – Maps, timelines, and more</li> <li>○ Provide personalized information services</li> <li>○ Support online research work (e-research)</li> </ul>
<p style="text-align: center;"><b>BASIC</b></p> <ul style="list-style-type: none"> <li>○ Implement a good overview and navigation of resources</li> <li>○ Ensure richness and added value of information</li> <li>○ Provide effective data search and filter functionality</li> </ul>	<p style="text-align: center;"><b>SUPPORT</b></p> <ul style="list-style-type: none"> <li>○ Help users understand and use specific terminology</li> <li>○ Provide data preview and license information</li> <li>○ Support different download options and external applications</li> </ul>

## 6.2 Specific recommendations

In this section, we specify the above mentioned high-level recommendations in more detail, by listing specific recommendations for the overall approach, design, services and features of the ARIADNE portal that were made by lead users.

**Implement a good overview and navigation of resources – recommendations:**

- Design a highly functional as well as attractive portal, e.g. with regard to an overview of searchable content/data and portal navigation.
- Provide statistical information on the quantity and distribution of the data (e.g. per type of data, provider, country/area, period, etc.).
- Enable a good understanding of the data resources and how they can be searched, accessed and (re-)used.

**Ensure richness and added value of information – recommendations:**

- Seek to integrate many resources from affiliated institutions as well as from other initiatives within and beyond the archaeology sector (e.g. remote sensing or molecular biology data).
- Emphasise the European/international dimension of the initial ARIADNE resources.
- Create added value through linking to data and publication resources not held within the ARIADNE Registry.
- Actively promote a culture of open data sharing in the sector, through making users aware of the importance of sharing ("give and take") and point to good practice guides and community repositories (including archives in the ARIADNE federation).

**Help users understand and use specific terminology – recommendations:**

- Provide aids for users not familiar with specific terminology/vocabularies (e.g. look up or dynamically invoked glossary or scope notes of thesauri).
- Consider terminology support within search features as well as information pages for data resources, where possible in multi-lingual form.
- Promote the provision of terminologies according to standards (e.g. ISO 25964) and in (semantic) web-based formats as required for machines and humans.

**Integrate and link information resources – recommendations:**

- Make users aware that the portal does not store or create databases of primary data but operates based on the metadata of, and semantic relations between, data/content collections and items.
- Provide integrated access to data and publications (i.e. include metadata of document archives and publishers).
- Specifically support the inclusion and linking of information (metadata) from archaeological grey literature, which may be produced through Natural Language Processing (NLP).

**Follow and promote Linked Data principles – recommendations:**

- Deploy Linked Open Data (LOD) to integrate information within the portal and to link to external resources which follow LOD principles (e.g. HTTP URLs).
- Demonstrate advantages of Linked Data to encourage further uptake of LOD principles by archaeological institutions and projects.
- Provide an LOD triple-store so that also external application developers can exploit resources of the portal for added value services (e.g. interlinking of databases).

**Provide effective data search and filter functionality – recommendations:**

- Investigate the adequate set and implementation of search options directly with members of the user community (various options will be suggested which should be scrutinized).
- Seek regular feedback on implemented solutions by the wider community.

**Visualize data resources: Maps, timelines, and more – recommendations:**

- Offer search & filter functionality based on maps as well as date-ranges (timelines) & named periods. Both features will be required to allow users to select potentially relevant datasets.
- Consider also visualizations based on Linked Data (e.g. density and web of links of the Linked Data graph around concepts).

**Provide data preview and license information – recommendations:**

- Implement a data preview mechanism that enables portal users to check if discovered data resources are actually relevant for the study purpose.
- Make users aware of available (or missing) license information of the data provider, e.g. to identify restrictions which impede re-use.
- Enable portal users to filter available data according to allowed usage, from Public Domain to fully (c)-restricted.

**Support different download options and external applications – recommendations:**

- Support download of data/metadata, including single item and bulk download as well as export/download in different open formats (for data that is not directly accessible, refer to the download page of the data repository).
- Provide interfaces to allow external applications to exploit available data, metadata and conceptual knowledge resources (e.g. a well-documented API, OAI-PMH target, SPARQL endpoint).
- Consider also RSS feeds for researchers and websites of research communities in particular subjects or geographic regions (e.g. alerts on new datasets in digital archives).

**Provide personalized information services – recommendations:**

- Implement personalized information services which meet clearly evaluated needs of large segments of users.
- Prioritise features which support users' control of specific searches, results and use of data.
- Consider notifications (alerts), e.g. e-mail lists or RSS feeds, as good candidates for personalized services.

**Enable linking and exchange of professional information – recommendations:**

- Enable linking of existing profiles of researchers on institutional websites and professional networking platforms to portal web pages (e.g. "MyData" pages).
- Investigate how the portal might enable effective expert discussion.

**Support online research work (e-research) – recommendations:**

- Support integrated access as required for studying various research resources online (e.g. linking and comparing content).
- Provide or link to tools which enable researchers to extract and combine data (e.g. images from different databases, numeric data to produce a derived dataset).
- Provide or link to tools for data processing and analysis (e.g. statistical analysis, image data processing and analysis).



### 6.3 General conclusions for the ARIADNE e-infrastructure and service development

ARIADNE is the main EU FP7 Integrating Activity project in the field of archaeology. The project addresses the fragmentation of archaeological datasets and limited online access to openly shared data in Europe. Therefore the project is developing an e-infrastructure that will allow for interoperability of existing and newly built digital archives and, based on this interoperability, cross-archive search, access and re-use of available data.

The creation of such an e-infrastructure will be a considerable step forward in the archaeological domain. The e-infrastructure will provide a common space where the currently dispersed resources can be uniformly described in the ARIADNE Registry and searched and accessed by the research community and other user groups on the Data Portal.

With regard to the development of the Data Portal, some general conclusion can be drawn from the recommendations of the lead user survey:

- Conclusion 1:** The recommendations confirm the overall focus of the ARIADNE project on data discovery and access services.
- Conclusion 2:** The prime attention of the design and interfaces of the data portal should be an overview of what data is accessible, including statistical information on quantity, types, distribution (e.g. country/area, period, etc.).
- Conclusion 3:** The portal should focus on the European/international dimension. Lack of underlying resources (per country, type of data, etc.) should not be seen as a deficit, but used to promote data mobilization (e.g. implementation of national data archives).
- Conclusion 4:** Added value should be created through also linking data and publication resources not held within the ARIADNE Registry (e.g. metadata of document archives and open access publishers).
- Conclusion 5:** Linked Open Data (LOD) can play a core role for value generation, but further uptake of LOD principles by archaeological institutions and projects must be encouraged.
- Conclusion 6:** In the development of the data search, access and other services members of the user community must be thoroughly involved and regular feedback on implemented solutions sought by the wider community.
- Conclusion 7:** User-focused development of the portal services and applications (relevance, usability, user-friendliness) should be at the top of the project's priorities.
- Conclusion 8:** Services for websites of research communities in particular subjects or geographic regions (e.g. alerts on relevant datasets) could greatly expand the reach of the data portal and, in turn, promote further data mobilization.
- Conclusion 9:** Full exploitation of the data resources (incl. metadata, conceptual knowledge) should be enabled by interfaces for external applications (e.g. a well-documented API, OAI-PMH target, SPARQL endpoint).
- Conclusion 10:** Support of e-research/science should in the first place be provided through integrating access to data resources and pointing users to existing tools for data extraction, processing and analysis.

## Annex I: Questionnaire & guidelines provided for the portal evaluation

This Annex contains the instructions for the lead user panel that evaluated the research archives and portals, and the questionnaire guideline.

### Introduction and general guidelines

---

Dear Colleague,

Thank you for participating in this lead-user survey about research-related online portals. The survey is carried out in the context of the ARIADNE project ("Advanced Research Infrastructure for Archaeological Dataset Networking in Europe") which is funded under the European Community's Seventh Framework Programme (see: <http://www.ariadne-infrastructure.eu>).

The goal of this survey is to learn more about current practices and requirements of archaeological researchers in how they use online portals and archives.

The survey is structured in three modules.

- **Module A** is about your **research activities in general**, and your **use of archaeological data**, irrespectively of the portal or archive you will evaluate.
- **Module B** contains the **questions about the specific portal** you have selected for the evaluation; Annex I provides a list of portals, with a suggestion for each partner.
- **Module C** is the space for your **suggestions and ideas** how data portals could/should be improved in the future.

To participate in the survey, we recommend you to proceed as follows:

1. **Complete Module A.** You can directly start with answering the questions in Module A without first checking the portal you will evaluate.
2. **Select a portal.** You can choose the portal suggested in Annex I, or propose another portal for the evaluation which you are using in your work and are familiar with. If you go for the latter option, please inform us beforehand about your choice.
3. **Familiarise yourself with the selected portal.** If you are not already familiar with the portal, we recommend that you do some preparatory work before answering the questions of Module B. Please check out the portal (e.g. the scope of its offer, the features, the way it works) before you start with the actual evaluation.
4. **Specify a "dummy" research task.** To assess the offer and the usability of the portal, we suggest that you define beforehand a specific task which you want to accomplish on the portal. For instance, you could specify a 'realistic' research task which could be part of a research project you are involved in. This task will typically consist in searching for specific data/datasets, downloading the available data, and checking the quality and completeness of the obtained data.
5. **Carry out the task.** Try to complete the task you have specified on the portal.
6. **Complete Module B.** Please consider Module B as the "evaluation report" of this survey. We ask you to provide feed-back on the portal (what is good, what could be improved, what was surprising), using the questions in the module as a checklist and guideline. Feel free to go beyond the listed questions in the feed-back you are providing.
7. **Complete Module C.** Finally, please provide ideas and suggestions how online portals for researchers could be improved in the future. Feel free to include both suggestions for simple improvements in details, as well as "crazy" (visionary) ideas which are unlikely to be realised

in the near future. Again, use the questions as a guideline, but feel free to make suggestions beyond the proposed issues.

Preferably, please use this document as the questionnaire, fill in your answers directly into the document and send the completed document back to us. You need not bother about formatting issues. If, for technical or other reasons, you prefer to use another document (or empty document) for writing down your answers, you can do so as well. In this case, please make sure to include as references the question numbers.

Please note that the size of the text-boxes for your answer is not indicative for the length of the answers we are expecting! In some cases, your feedback may be only a line or two; in other cases, you may want to provide more extensive information and insert text that runs over a page. You are totally flexible in this regard.

### Helpdesk

If there are any uncertainties in how to carry out the evaluation, feel free to contact us (by e-mail or telephone) any time. Contact persons are:

Hannes Selhofer

Tel. +43.662.2288-254

e-Mail: [hannes.selhofer@salzburgresearch.at](mailto:hannes.selhofer@salzburgresearch.at)

Guntram Geser

Tel. +43.662.2288-303

e-Mail: [guntram.geser@salzburgresearch.at](mailto:guntram.geser@salzburgresearch.at)

We highly appreciate your taking the time for this assessment and are looking forward to receiving your evaluation reports.

Best regards



Hannes Selhofer

for the ARIADNE WP2 research team

## Module A) Background and context: your research activities and general data requirements

The questions in this module concern your research activities in general, and your use of archaeological data, irrespectively of the portal you will evaluate. The information you provide us in this section will help to put your assessment and suggestions into perspective, and to better understand the general work-flows of data usage in archaeological research projects.

### A.1) Basic information

*Please describe briefly your research organisation (if you are affiliated with an organisation) and your current position in the organisation.*

<b>Organisation:</b>	(Name)
<b>Website:</b>	
<b>Your department / research group:</b>	
<b>Your position:</b>	

**Short description of your research group**  
(Please describe)

*How many years have you been involved in archaeological research, and for how many years (roughly) have you been using online databases / portals in this context?*

<b>Professional experience in archaeology</b>	___ Years
<b>Use of online databases / portals</b>	___ Years

*Please describe briefly your own research focus (e.g. in terms of research domains, themes, geographic dimension).*

**Your research focus**  
(Please describe)

*Which kind of data are you mainly using in your research? Please describe the type of data (e.g. excavation data, field survey data, laboratory data...)?*

**The data you are mainly using in your research**  
(Please describe)

*What are the main sources from which you normally obtain your data (such as: institutional databases, records of governmental organisations, own or external laboratory,...)?*

**The main sources of the data you are using**  
(Please describe)

*What are the main online sources from which you obtain research data? Please name the 3-5 most important archives, repositories or portals which you are using in your research, and describe which type of data you collect from these sources. – If there are no specific archives you are using, please describe the type of online sources where you are searching for data / information. How important are these sources compared to the general sources you mentioned in question A.4?*

No.	Name of the archive / repository	Data obtained
1	(Name & website – if any)	(Please describe)
2		
3		
4		
5		

**Other main online sources for relevant data / information (apart from the ones listed above)**  
(Please describe)

*What are the main shortcomings and problems you have experienced in the use of online archives and databases in your own work? (Think, for example, about issues such as quality of description (metadata), access to datasets, relevance and completeness of data, cost issues, licensing)*

**Main shortcomings and problems related to the use of online archives**  
(Please describe)

*Please outline your typical work flow in an archaeological research project with regard to the use of online archives and portals. In which stage of the project would you perform which activities?<sup>24</sup>*

Stage	Activities (with regard to use of archives / portals)
(Describe Stage 1)	•
(Describe Stage 2)	•
(Describe Stage 3)	•

*To what extent and in what way is data which your research group is producing typically being published (i.e. made available to a certain community beyond your own institute)?*

**Typical ways of publishing data after completion of a project**  
(Please describe)

<sup>24</sup> Stages might include the preparatory and proposal phase (before the project start, including possibly the preparation of a research proposal for funding), the set-up phase of a project (planning, preparatory work), the main research stage, the analysis phase and, finally, the publication and archiving of results. This is just indicative, feel free to select other stages.

Are you (at least occasionally) depositing research data from your research projects in an online database (other than your internal institutional database) which can be accessed by other researchers? If yes, please name the most important databases / repositories where you have deposited research data.

No.	Name of the archive / repository	Focus of the archive	Website
1			
2			
3			

Any other feedback: Is there any other information / feedback you would like to provide on your research activities, as far as relevant for this survey, which was not covered by the above questions?

**Other information relevant for this survey**  
(Please describe)

### Module B) Assessment of a specific archive / portal

The questions in this module are about the specific archive / portal you have selected for the evaluation. They concern the available features, an assessment of their usability, good practices that could serve as example for other archives, barriers or problems you are experiencing when using the archive.

#### B.1) The portal and your approach for the evaluation

Which portal have you selected for the evaluation? Please provide some basic information about the portal you are going to evaluate.

Name	
Portal / website (URL)	
Provider	
Domain / thematic focus	
Geographic focus	
Registration necessary?	Is a registration necessary to use the services (or to use some of the services)?

Please describe briefly the type of data offered and to what extent is or could be relevant for your own research activities.

**Type of data offered**  
(Please describe)

**Relevance of the portal for your own research**  
(Please assess / describe)

*The approach and hypothetical research task: please describe briefly how you have planned to carry out the evaluation. Have you specified a hypothetical research assignment which you will carry out in order to explore the functionalities? If so, please briefly describe this assignment.*

#### Your approach for the evaluation

(If you have specified a hypothetical "research assignment" for carrying out the evaluation, please briefly describe this assignment)

#### The conceptual model for the evaluation

The following model provides a guideline or checklist for the aspects you could consider in your evaluation. The conceptual model which we propose as a baseline for the evaluation (see Figure below) consists of two groups of functions:

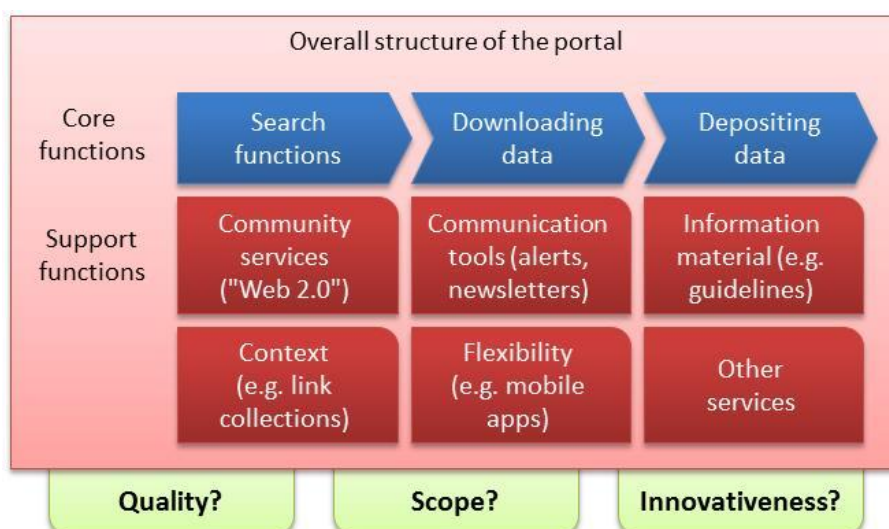
- **Core functions** of a data portal (in our view, these are: search functions, functions for downloading data, and functions for depositing one's own research data);
- **Support functions**, such as add-on information services (e.g. guidelines, tools, news services), or communication services.

Not all of these functions, in particular the support functions, will be available on all portals. If a function is not existent in a portal, please say so.

For each of the available functions we would ask you to perform a short assessment and provide feedback about your experience. The assessment could focus on the following dimensions:

- **Scope:** What are the functionalities that are offered? What are you missing?
- **Quality:** The usability of the various functionalities. Please highlight specifically useful features and describe what makes them useful for you.
- **Innovativeness:** Is there anything that was new to you / which you have not seen offered in this way elsewhere, and which could be a model for other services of this type?

*Figure: Components of the evaluation: a conceptual model for the assessment*





**B.2) The overall structure and “organisation” of the portal**

*How would you assess the overall structure and organisation on a scale from 1 (very good) to 5 (poor)?*

1	Very good	<input type="radio"/>
2	Good	<input type="radio"/>
3	Average	<input type="radio"/>
4	Fair	<input type="radio"/>
5	Poor	<input type="radio"/>

*Please provide feedback on the overall design and structure of the portal. Is it well organised (why / why not), is it easy to navigate, is the amount of information displayed on each page adequate, ...?*

**Overall design and structure of the portal – your assessment**

(Please describe)

*Are there any specific elements in the overall structure which you consider a “good practice” that could serve as a model for other portals? If so, please describe.*

**Overall design and structure of the portal – good practices**

(Please describe)

**B.3) The search functions**

*How would you assess the search functions which the portal is offering on a scale from 1 (very good) to 5 (poor)?*

1	Very good	<input type="radio"/>
2	Good	<input type="radio"/>
3	Average	<input type="radio"/>
4	Fair	<input type="radio"/>
5	Poor	<input type="radio"/>

*Please provide feedback on the search functions of the portal. Are they well organised (why / why not), is it easy and convenient to find data, ...?*

**Search functions of the portal – your assessment**

(Please describe)

*Are there any specific elements in the search functions which you consider a “good practice” that could serve as a model for other portals? If so, please describe.*

**Overall design and structure of the portal – good practices**

(Please describe)

**B.4) The mechanisms for downloading data**

*How would you assess the download mechanisms which the portal is offering on a scale from 1 (very good) to 5 (poor)?*

1	Very good	<input type="radio"/>
2	Good	<input type="radio"/>
3	Average	<input type="radio"/>
4	Fair	<input type="radio"/>
5	Poor	<input type="radio"/>

*Please provide feedback on the download mechanisms of the portal. Are they well organised (why / why not), is it easy and convenient to download data, ...?*

**Download mechanisms of the portal – your assessment**

(Please describe)

*Are there any specific elements in the download mechanisms which you consider a “good practice” that could serve as a model for other portals? If so, please describe.*

**Download mechanisms of the portal – good practices**

(Please describe)

**B.5) Options for uploading / depositing and sharing own data**

*Does the portal provide the opportunity to users to upload / deposit and share own data?*

Yes	<input type="radio"/>
No	<input type="radio"/>

If “no”, the following questions of B.5 are not relevant.

*How would you assess the uploading / depositing and sharing mechanisms which the portal is offering on a scale from 1 (very good) to 5 (poor)?*

1	Very good	<input type="radio"/>
2	Good	<input type="radio"/>
3	Average	<input type="radio"/>
4	Fair	<input type="radio"/>
5	Poor	<input type="radio"/>

*Please provide feedback on the uploading / depositing and sharing mechanisms of the portal. Are they well organised (why / why not), is it easy and convenient to deposit data, ...?*

**Depositing mechanisms of the portal – your assessment**

(Please describe)

*Are there any specific elements in the uploading / depositing and sharing mechanisms which you consider a “good practice” that could serve as a model for other portals? If so, please describe.*

**Depositing mechanisms of the portal – good practices**

(Please describe)

**B.6) Support functions offered by the portal**

*What are other functions / services which the portal offers in addition to providing access to research data? Please describe them briefly; some possible services are already mentioned in the table; feel free to extend the list with further services.*

Type of support services	Brief description
Community services	
Communication tools	...
Flexibility tools (e.g. mobile versions, apps)	...
Alerting / news feeds	
Information material (e.g. guides, thematic collections)	...
	...

*How would you assess the overall richness and usefulness of the support functions / services the portal is offering on a scale from 1 (very good) to 5 (poor)?*

1	Very good	<input type="radio"/>
2	Good	<input type="radio"/>
3	Average	<input type="radio"/>
4	Fair	<input type="radio"/>
5	Poor	<input type="radio"/>

*Please provide feedback on the support services of the portal. For instance: How are they organised? What is particularly useful? Is there a vivid user community exchanging views?*

**Support functions / services offered by the portal – your assessment**

(Please describe)

*Are there any specific services which you consider a “good practice” that could serve as a model for other portals? If so, please describe.*

**Support functions / services offered by the portal – good practices**

(Please describe)

## Module C) Ideas and suggestions

The questions in this module are about ideas for important and innovative features and functions which you would like to see in online archives and portals for researchers. The questions go beyond the specific portal which you have evaluated, but may draw from this assessment (e.g. you may describe here something you missed from the portal). Feel free to include both suggestions for simple improvements in details, as well as “crazy” (visionary) ideas which are unlikely to be realised in the near future.

*What are the 3-5 features which you expect (or, if not yet existing, would like to get in the future) from an online portal which would be most helpful for your research activities? These can be specific services or tools – anything that makes a portal useful for you.*

#	Feature	Benefit for your work
1	Please describe the suggested feature	Please describe briefly why this would be important / helpful
2		
3		
4		
5		

*What is your general experience with search and other services on portals? What are good examples – what could be improved? What are the main bottlenecks currently where you would hope to see improvements?*

**Experience with regard to search and other services on portals**  
(Please describe)

*Are there any examples of useful services from portals other than the one you have evaluated which you would like to recommend?*

**Good practice examples from other portals**  
(Please describe)

*Any other suggestions and ideas: if you have any further suggestions, ideas or recommendations with regard to research portals, please share them with us!*

**Suggestions, ideas and recommendations**  
(Please describe)

## Annex II: Overview of portals and evaluators

	Digital archive/portal	Website	Evaluated by (organisation)*
General archaeology portals (see Section 4.2)	ADS	<a href="http://archaeologydataservice.ac.uk">http://archaeologydataservice.ac.uk</a>	DAI
	Arachne	<a href="http://arachne.uni-koeln.de/">http://arachne.uni-koeln.de/</a>	Athena Research DAI MiBAC-ICCU
	DANS EASY (Data Archiving and Networked Services)	<a href="https://easy.dans.knaw.nl/ui/home">https://easy.dans.knaw.nl/ui/home</a>	DAI ARHEOVEST
	Fasti Online	<a href="http://www.fastionline.org/">http://www.fastionline.org/</a>	NIAM-BAS
	MAPPA	<a href="http://mappaproject.arch.unipi.it/?lang=en">http://mappaproject.arch.unipi.it/?lang=en</a>	PIN
	OpenContext	<a href="http://opencontext.org">http://opencontext.org</a>	OeAW/OREA
	Pleiades/Pelagios	<a href="http://pleiades.stoa.org/">http://pleiades.stoa.org/</a>	Athena Research AIAC
	Portal to the Past	<a href="http://www.portal.cig-icg.gr/">http://www.portal.cig-icg.gr/</a>	PIN
	tDAR – The Digital Archaeological Record	<a href="https://www.tdar.org/">https://www.tdar.org/</a>	DANS
Specialist archaeology portals (see Section 4.3)	CHARISMA Art Database Portal	<a href="http://archives-charisma-portal.eu/#home">http://archives-charisma-portal.eu/#home</a>	MNM-NOK
	Bone Commons	<a href="http://alexandriarchive.org/bonecommons/collections">http://alexandriarchive.org/bonecommons/collections</a>	The Cyprus Institute
	ceraDAT	<a href="http://www.ims.demokritos.gr/ceradat/?PHPSESSID=dnjxlxknI">http://www.ims.demokritos.gr/ceradat/?PHPSESSID=dnjxlxknI</a>	MNM-NOK
	CLAROS	<a href="http://www.clarosnet.org/XDB/ASP/clarosHome/index.html">http://www.clarosnet.org/XDB/ASP/clarosHome/index.html</a>	ADS
	Mapping Death	<a href="http://www.mappingdeathdb.ie">http://www.mappingdeathdb.ie</a>	MNM-NOK
Portals in other domains (see Section 4.4)	CIARD-RING	<a href="http://ring.ciard.net/">http://ring.ciard.net/</a>	Athena Research
	Eurostat	<a href="http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/">http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/</a>	Salzburg Research
	GBIF – The Global Biodiversity Information Facility	<a href="http://www.gbif.org/">http://www.gbif.org/</a>	Institute of Archaeology of the Czech Academy of Sciences, Prague
	GCMD – Global Change Master Directory	<a href="http://gcmd.nasa.gov">http://gcmd.nasa.gov</a>	The Discovery Programme (two evaluations – 12a, 12b)
	Morphbank: Biological Imaging	<a href="http://www.morphbank.net">http://www.morphbank.net</a>	ZRC SAZU
	PhytCore phytolith database	<a href="http://gepeg.org/enter_PCORE.html">http://gepeg.org/enter_PCORE.html</a>	MNM-NOK
	Scratchpads Biodiversity online	<a href="http://scratchpads.eu/">http://scratchpads.eu/</a>	Faculty of Archaeology, University of Leiden
	SEDAC - Socioeconomic Data and Applications Center	<a href="http://sedac.ciesin.columbia.edu/">http://sedac.ciesin.columbia.edu/</a>	Swedish National Data Service
Research comm.	Academia.edu	<a href="https://www.academia.edu">https://www.academia.edu</a>	Salzburg Research
	Mendeley	<a href="http://www.mendeley.com">http://www.mendeley.com</a>	Salzburg Research
	ResearchGate	<a href="http://www.researchgate.net">http://www.researchgate.net</a>	Salzburg Research

\*The evaluator was a researcher from the organisation listed who is experienced in the use of digital archives and was therefore selected for this task.