Making archaeological datasets F-A-I-R

A panoramic view of the ARIADNE story

Julian Richards

Archaeology Data Service, University of York

ARIADNE is funded by the European Commission's Seventh Framework Programme

Introducing the F-A-I-R principles

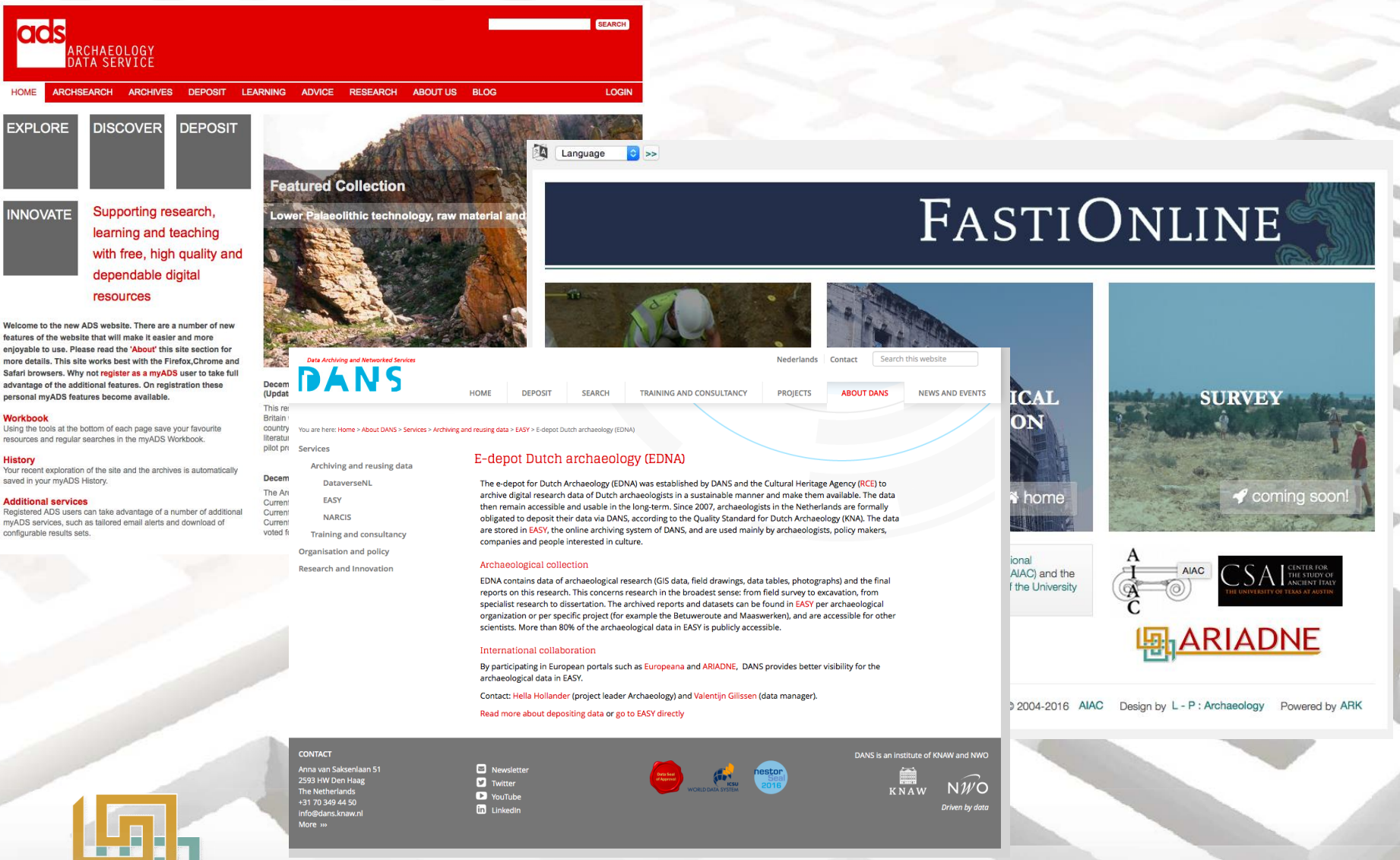
"One of the grand challenges of data-intensive science is to facilitate knowledge discovery by assisting humans and machines in their discovery of, access to, integration and analysis of scientific data."

- **Findable**
  - Data are described with good metadata
  - Metadata are indexed in a searchable resource
  - Data are assigned a Permanent identifier
- **Accessible**
  - Data should be open and online
- **Interoperable**
  - Use a formal, open, shared language for knowledge representation
- **Re-usable**
  - Data should have clear data licenses
  - Metadata should meet domain-relevant standards

Challenges and opportunities

- There is lots of data in archaeology
  - Fragmented, distributed, heterogeneous, hidden, short-lived
- IT is there to enable
  - Infrastructure, integration, standards, mapping, access, preservation

Life before ARIADNE…

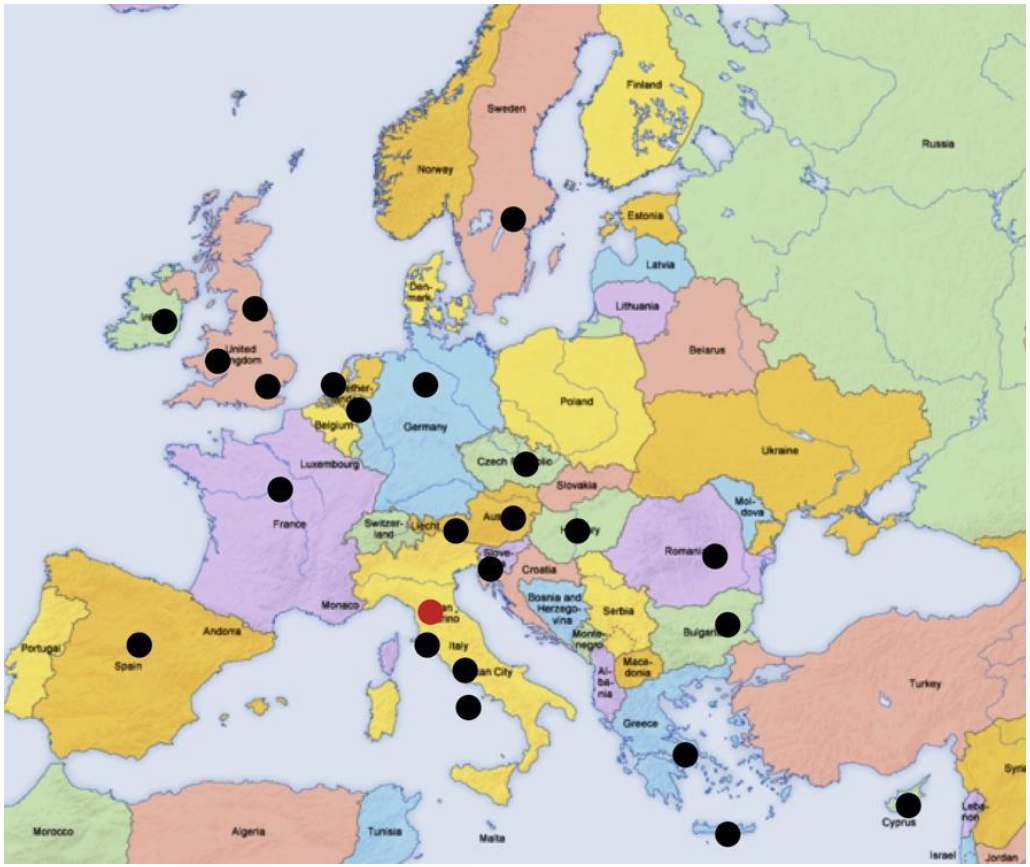


Project basics

- 4 year project (02/2013- 02/2017)
- FP7 Instrument “Integrating Activity”
- Funding 6.5m Euros
- Coordinators
  - Prof. Franco Niccolucci, University of Florence
  - Prof. Julian Richards, University of York
- Website: www.ariadne-infrastructure.eu

ARIADNE Community

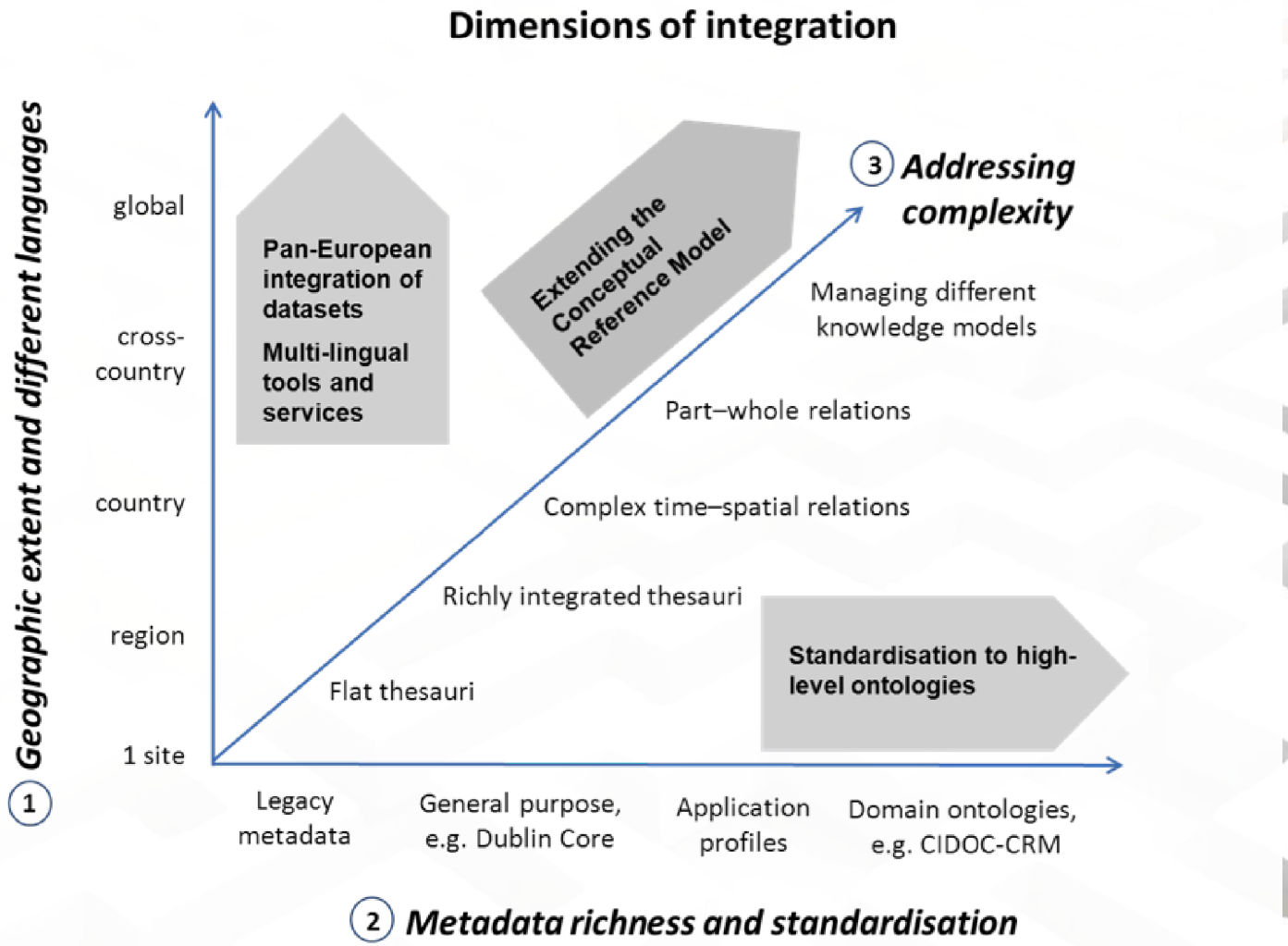
- 23 partners in 18 European countries
- 9 ICT organisations
- 14 archaeological organisations
- 15 Associate partners
- Community building
  - Transnational access
  - Training events
  - Special interest groups



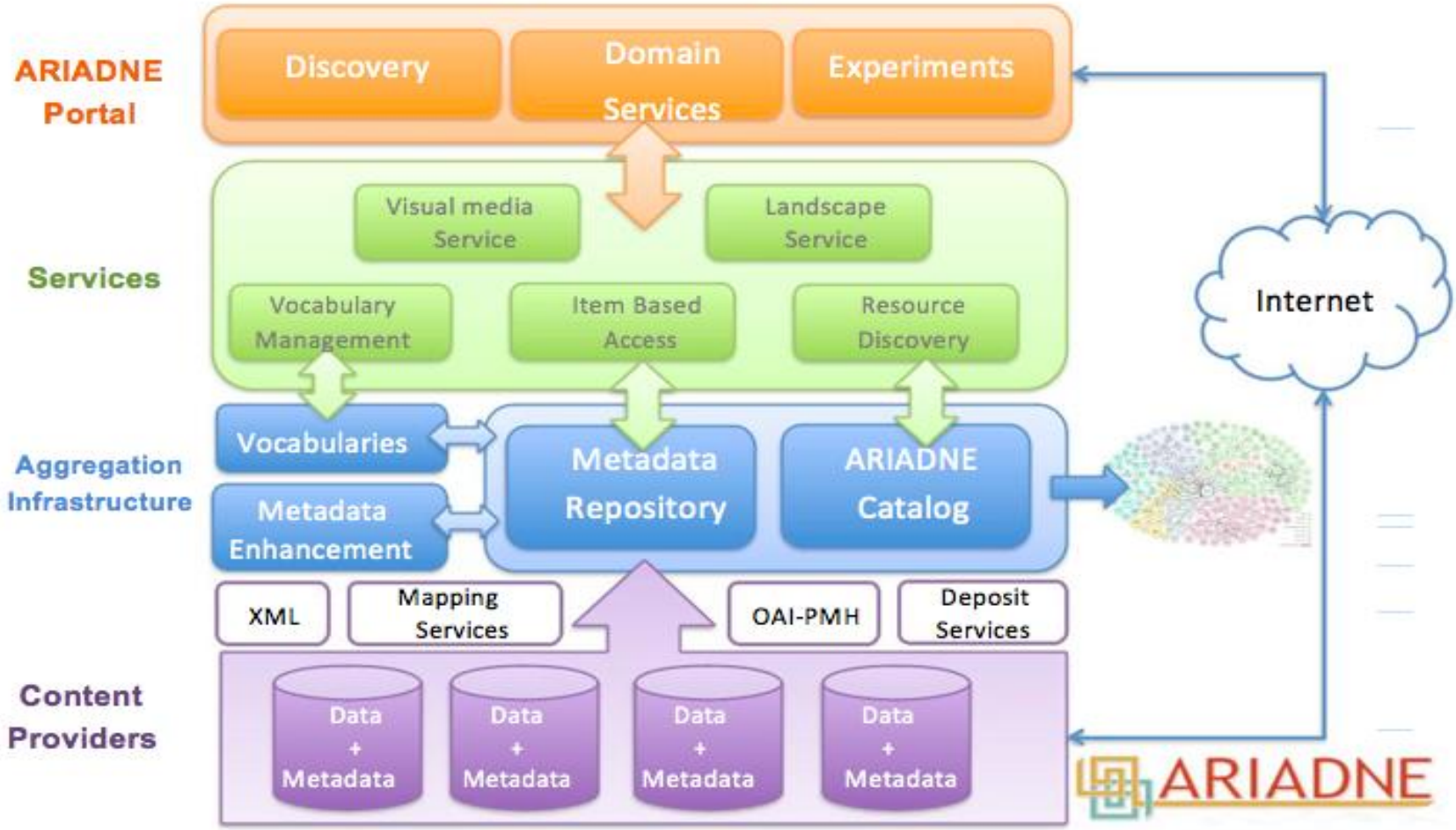
User Needs Research

- 94% of researchers agreed that it is important that datasets are available online in an uncomplicated way.
- 87% of researchers agreed that they often do not know what research data is available because it is stored in so many different places and databases.
- 74% of researchers consider it important to have easy access to international datasets.
- The additional effort required to prepare data (formatting, metadata, etc.) is considered to be a barrier to data sharing by 80% of researchers.
- The perceived lack of professional recognition and reward for sharing data is a barrier to data sharing for 72% of researchers.
- A lack of institutional or international repositories for archaeology data sets was a barrier to data sharing for 60% of researchers.

The ARIADNE roadmap

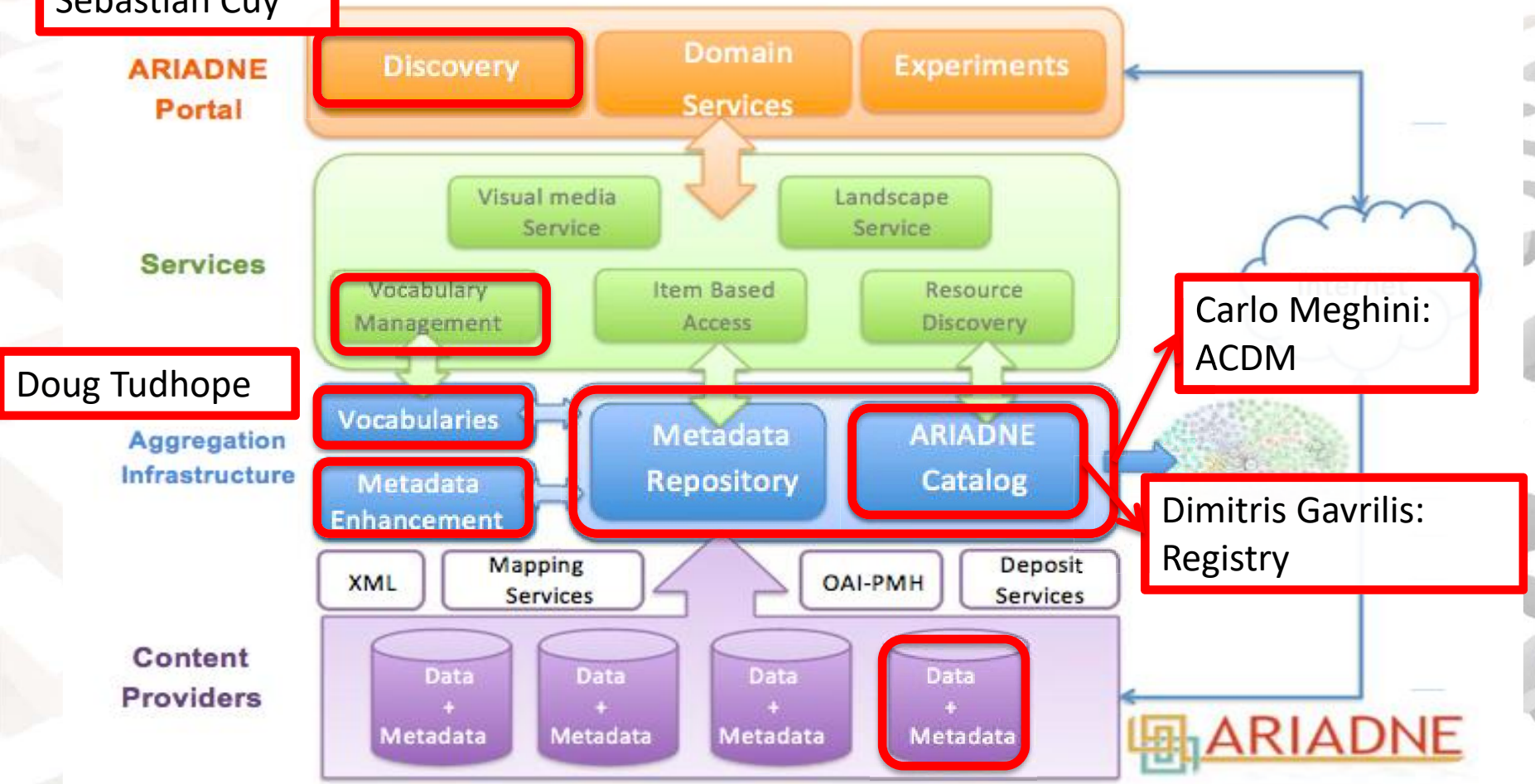


The ARIADNE Architecture



Improving findability and accessibility

Sebastian Cuy



Doug Tudhope

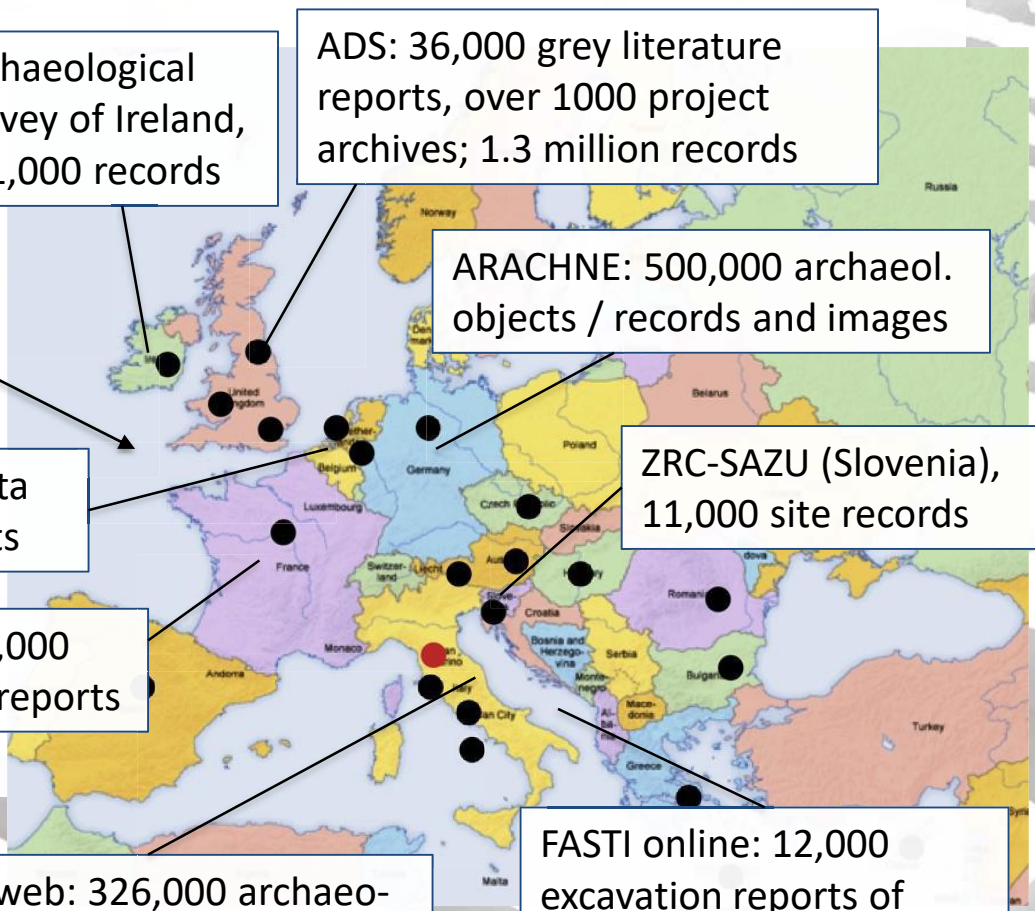
Carlo Meghini: ACDM

Dimitris Gavrilis: Registry

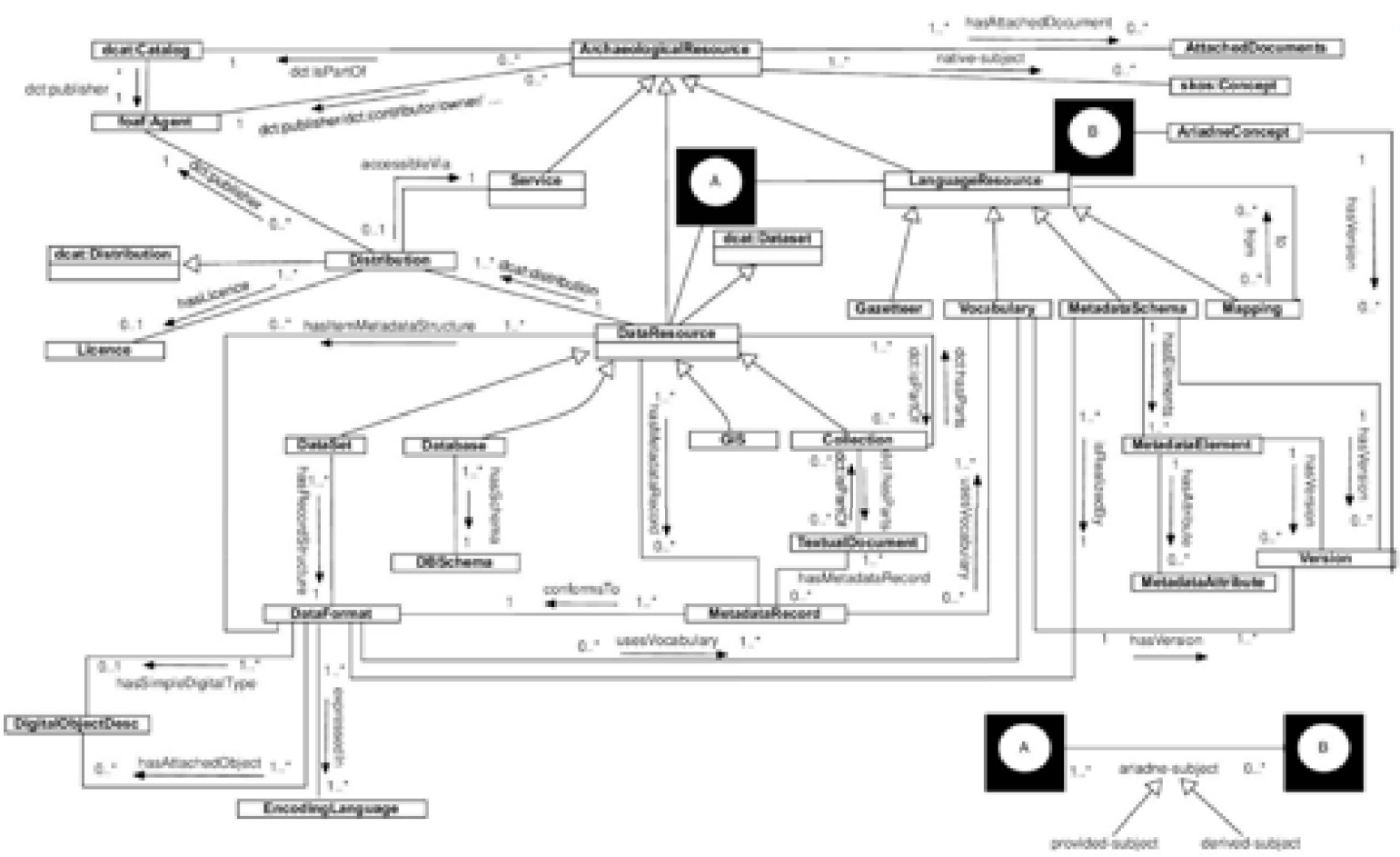
Achille Felicetti: TNA activity

Data sharing

- Mobilize
- Integrate
- Make accessible

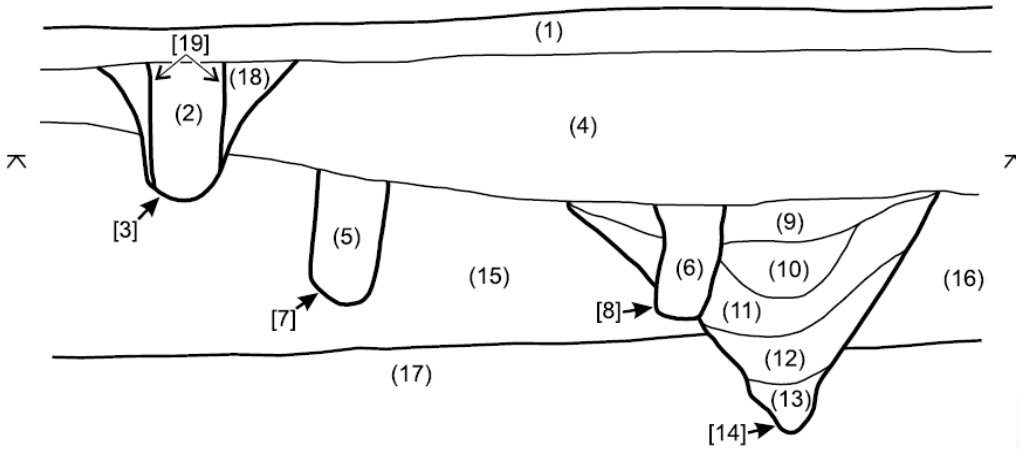


The ACDM: ARIADNE Common Data Model

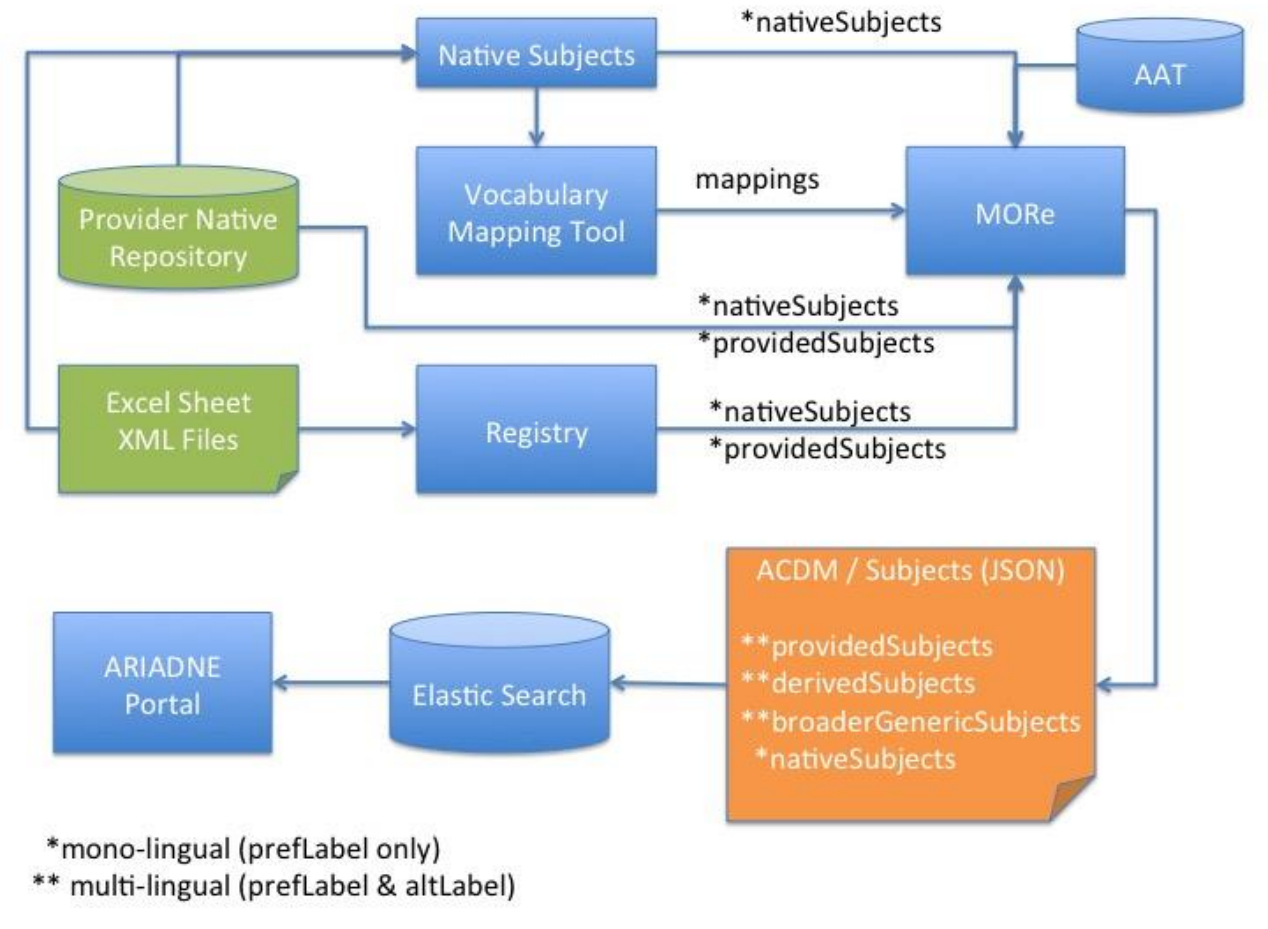


Achieving interoperability

- We have datasets in many languages and complying with many different standards
- ARIADNE uses the CIDOC CRM with extensions for archaeology to achieve integration
  - Existing datasets are mapped to the ARIADNE data model
  - Subject concepts are mapped to the Getty A&AT



Metadata enrichment in the Registry



Interoperability Framework

ARIADNE Users Framework

ARIADNE e-Infrastructure and Integrated Services

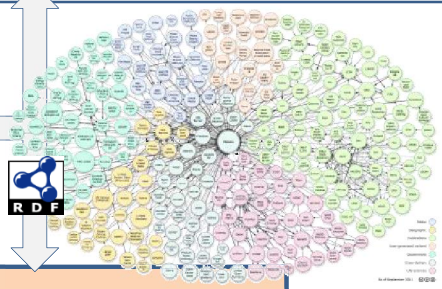
Other e-Infra & Services

L4: e-infra & integrated services

ARIADNE Portal
Cross-search and other applications

Search, browse, …

ARIADNE Registry
Datasets, metadata schema, KOSs, mappings,…, CRM



LOD cloud diagram, by R. Cyganiak & A. Jentzsch, http://lod-cloud.net

accData

Meta/data records => collect/transform2/RDF

collM collM collM

L3: Higher-level data managers

Data Centers
(nat./int., archaeology or general)

Subject/Domain-based Repositories & Portals
(nat./international)

Subject/Domain-based Referatories & Portals
(nat./international)

collM

ADS
eDNA
SND
…

depD+M

ARACHNE
Fasti Online
…

ARIADNE aggregators (nat./thematic)?

L2: Institutions, institutional repositories/DBs

Institutional Repositories / DBs

collM

depD+M

Several ARIADNE partners

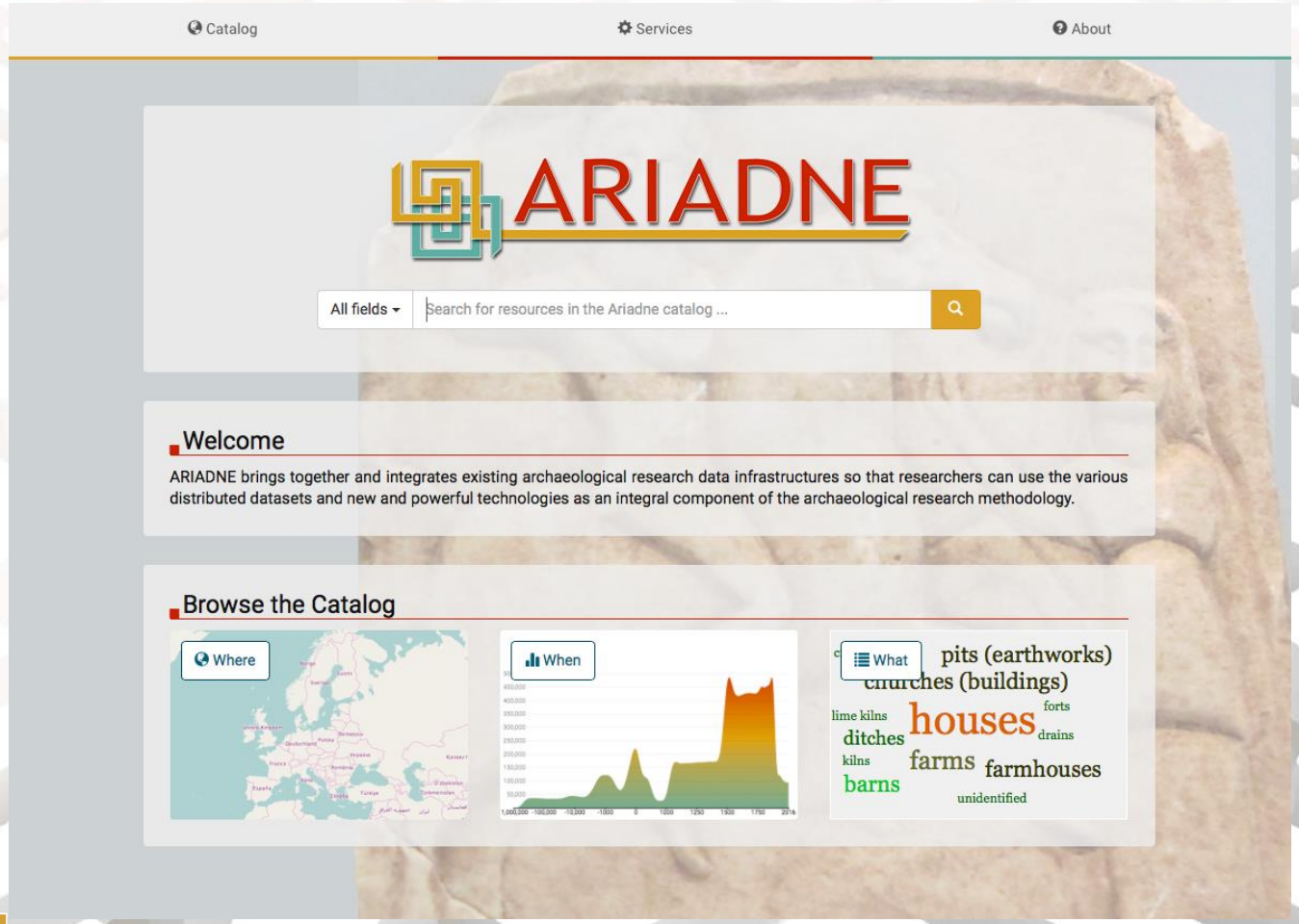
Other…

L1: Projects, research groups

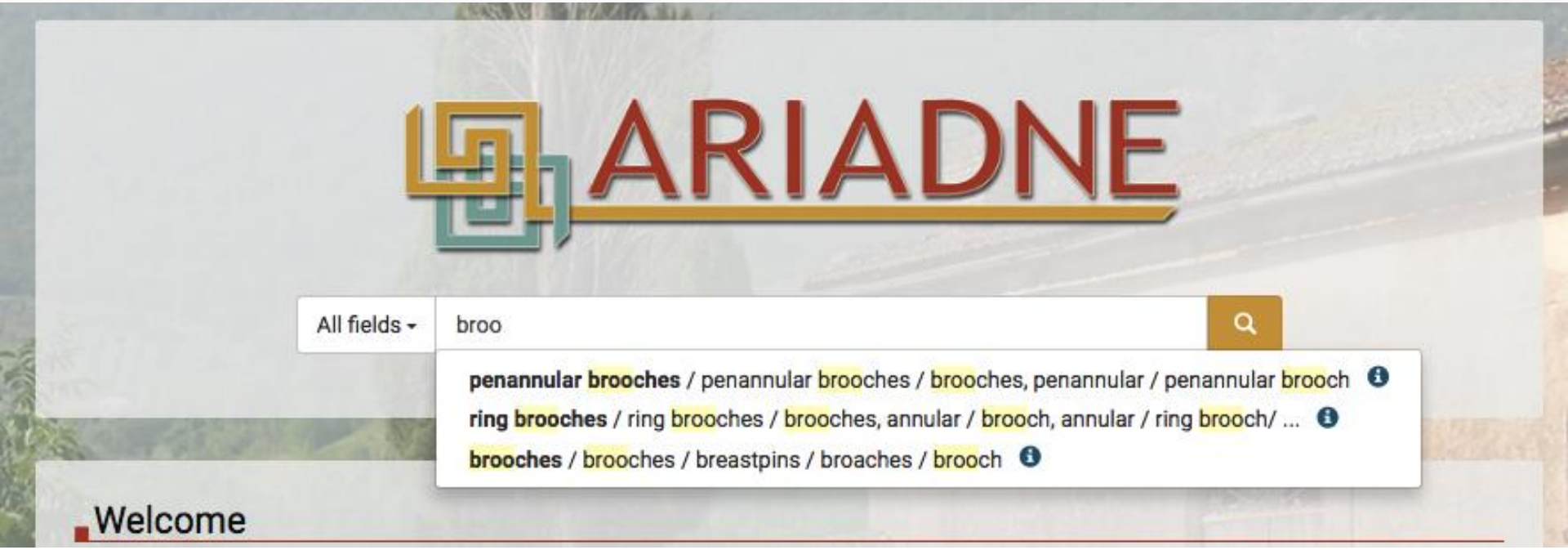
Research Projects …

Legend:
dep=deposit, coll=collect, acc=access;
D= Data, M=Metadata

The ARIADNE Portal

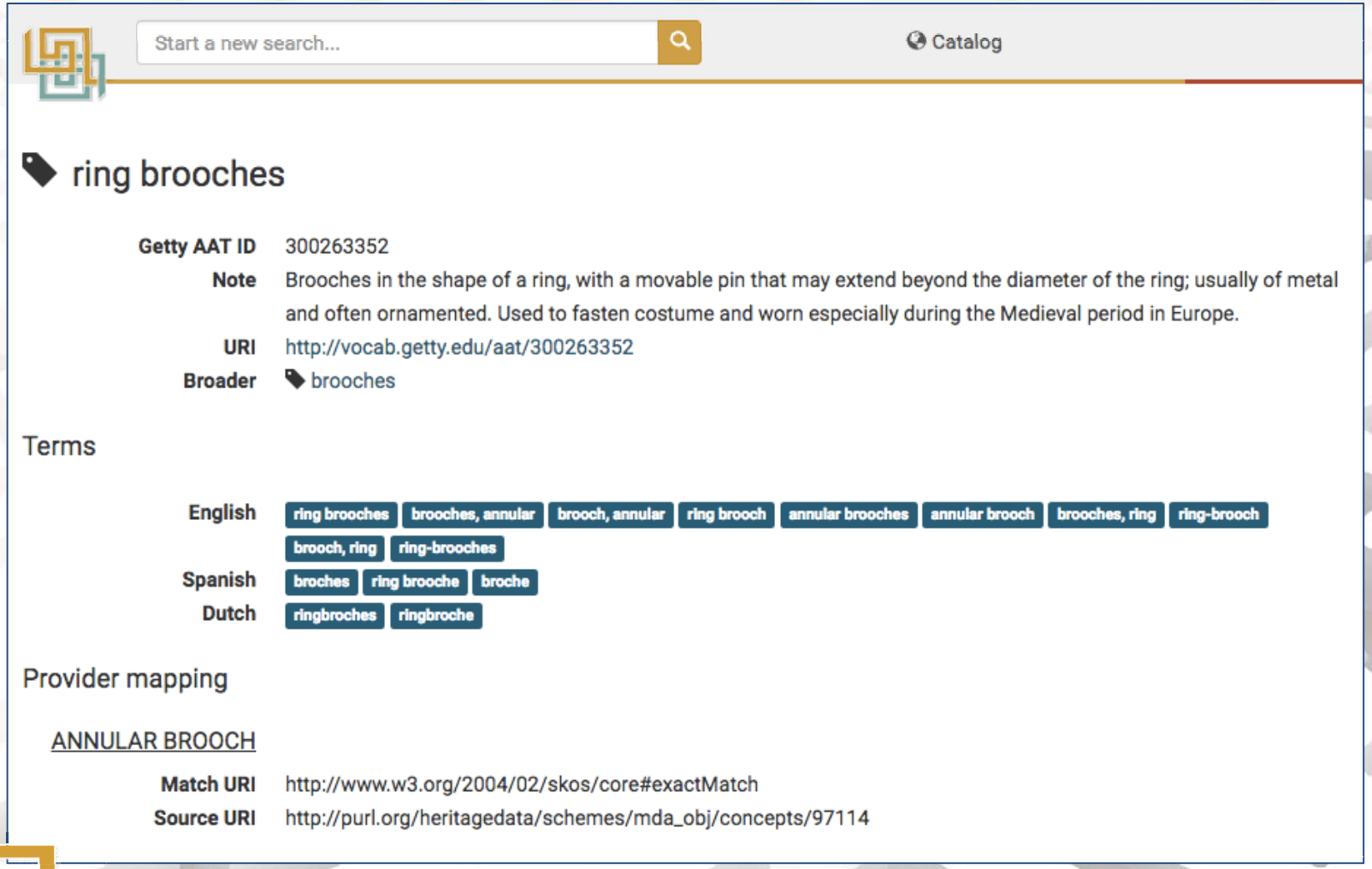


ARIADNE Portal

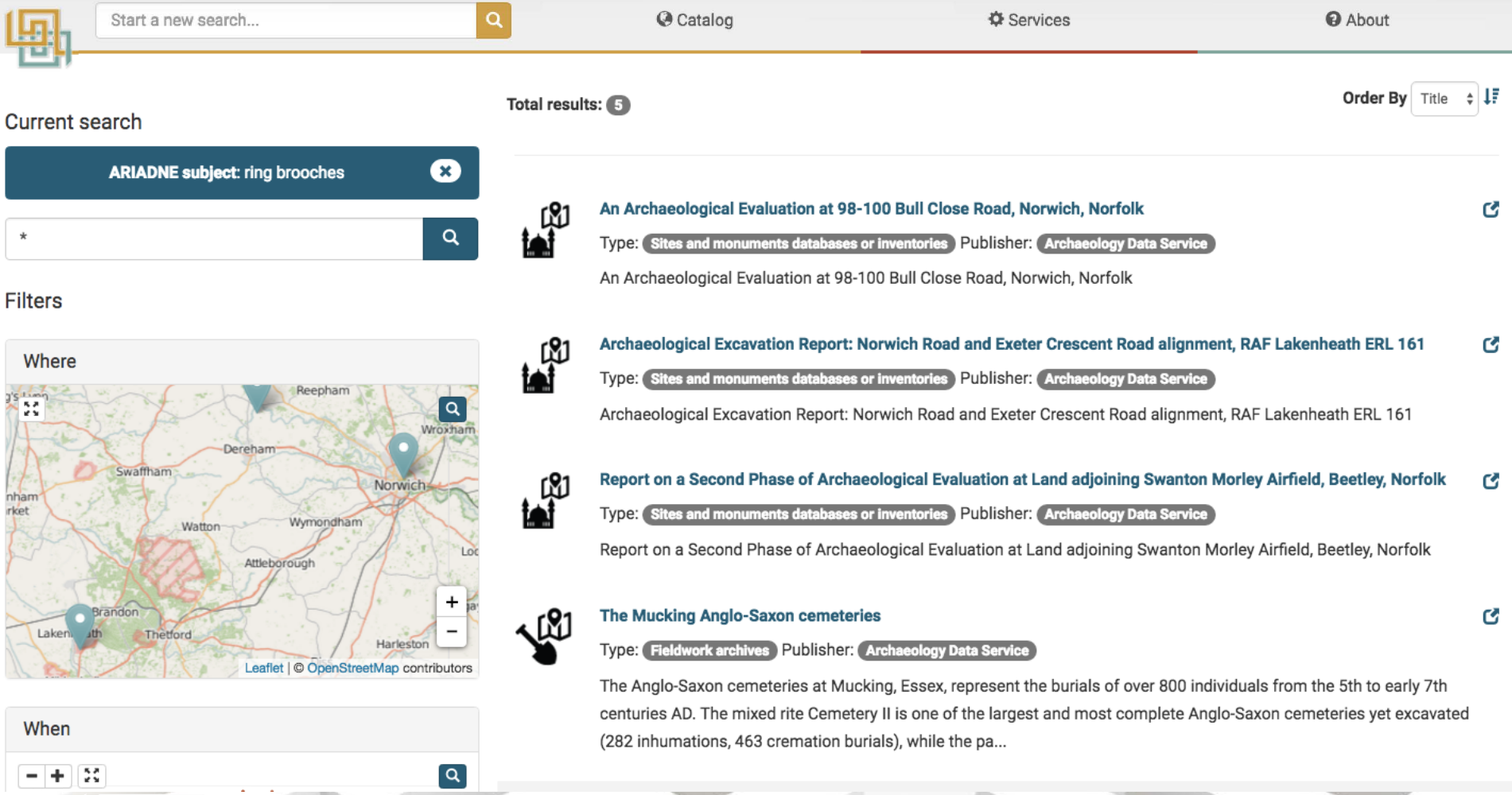


- Vocabulary concepts from the Getty Art & Architecture Thesaurus (AAT) are offered as users type a query
  - Enables query expansion
  - Supports multilingualism

Multilingual support



ARIADNE portal

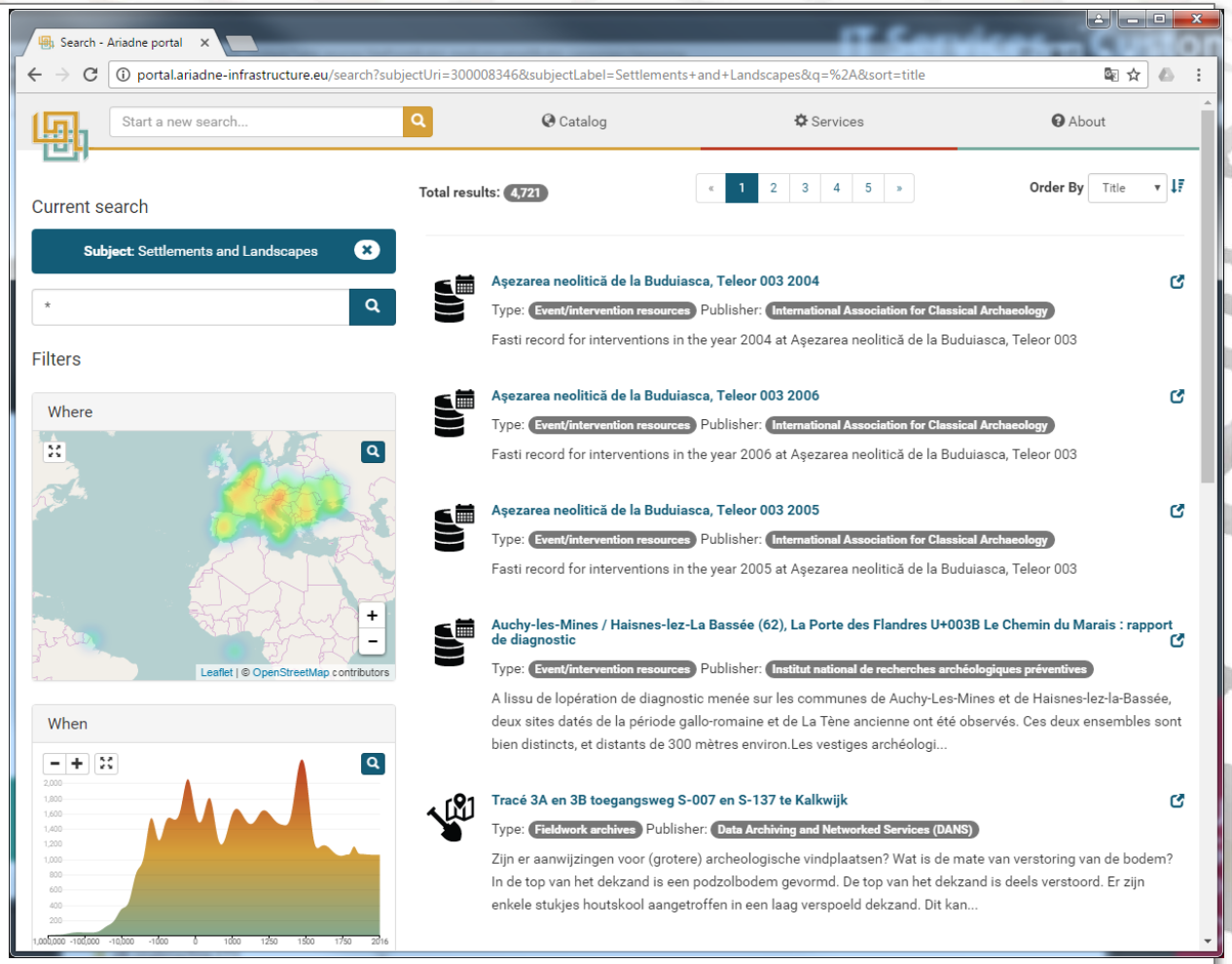


Multilingual results via AAT mappings

ARIADNE Portal
Query on AAT subject:
Settlements and Landscapes

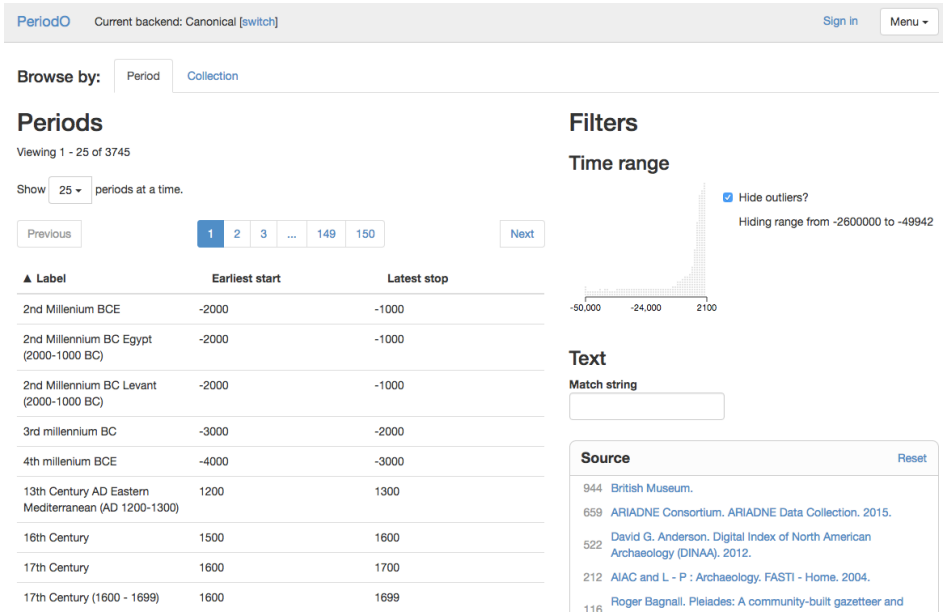
shows

results from IACA (Fasti), INRAP and DANS in multiple languages



ARIADNE and PeriodO

- PeriodO is a Linked Open Data gazetteer
  - Scholarly definitions of historical, art-historical and archaeological periods
  - Now includes period terms from ARIADNE partners
  - The terms have PeriodO URIs
- Allows for easier linking of datasets that define periods differently
  - (when is always linked with where in archaeology)



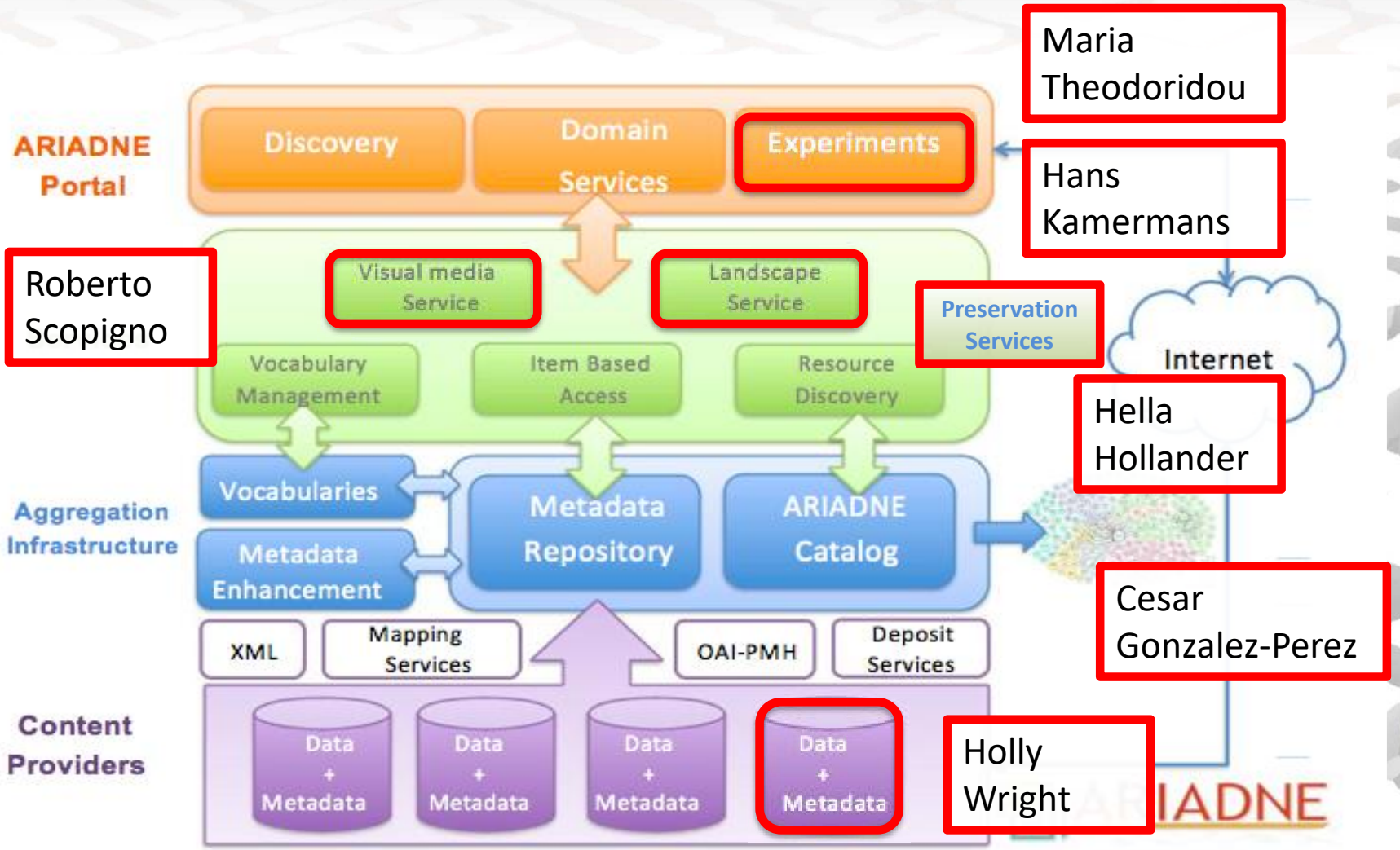
http://perio.do/

Transnational Activity

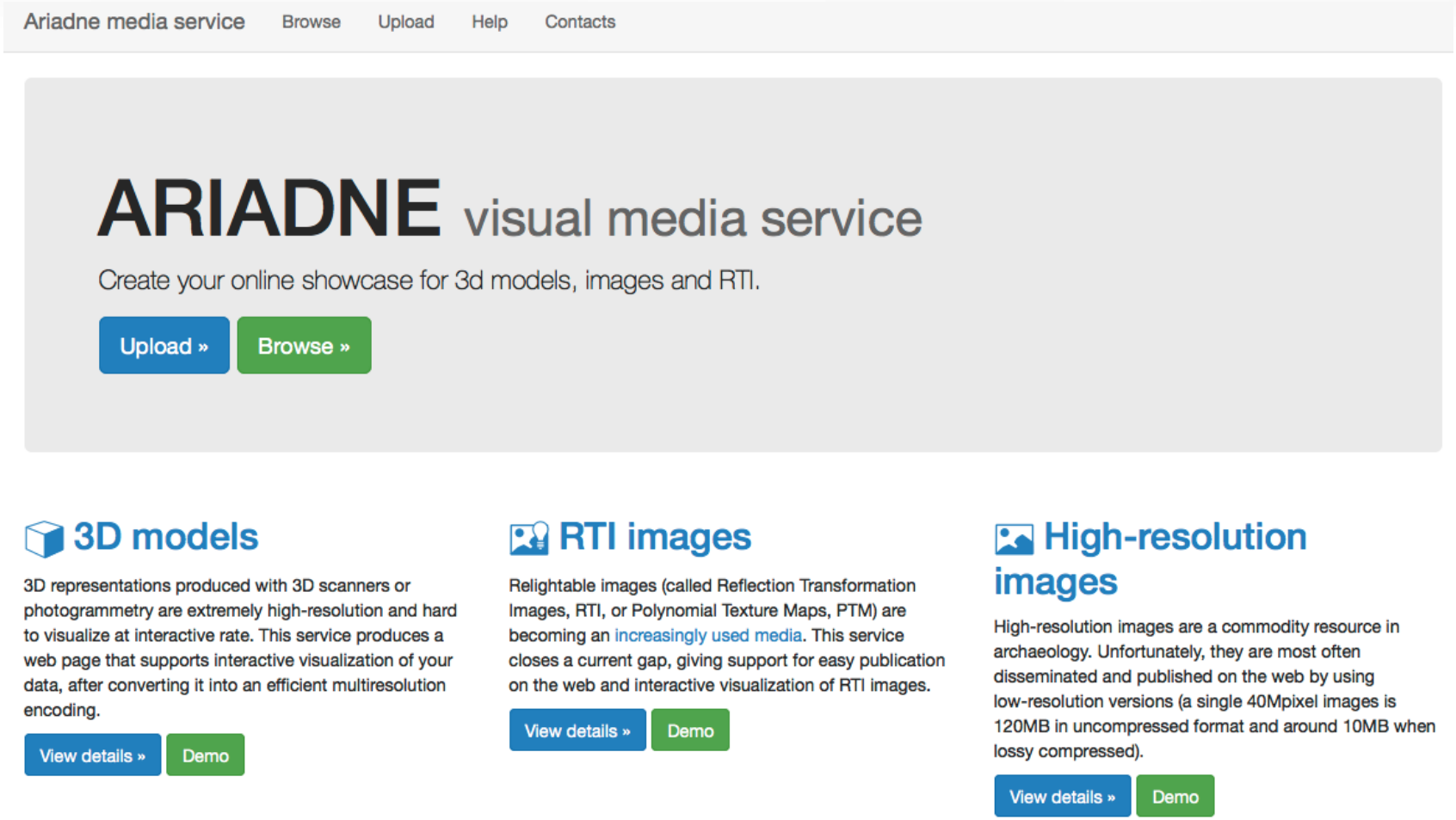
Training events

- 2D/3D documentation for archaeology
- Legacy data and dataset design
- Mapping existing datasets to CIDOC CRM

Improving interoperabilty and reusability



ARIADNE services

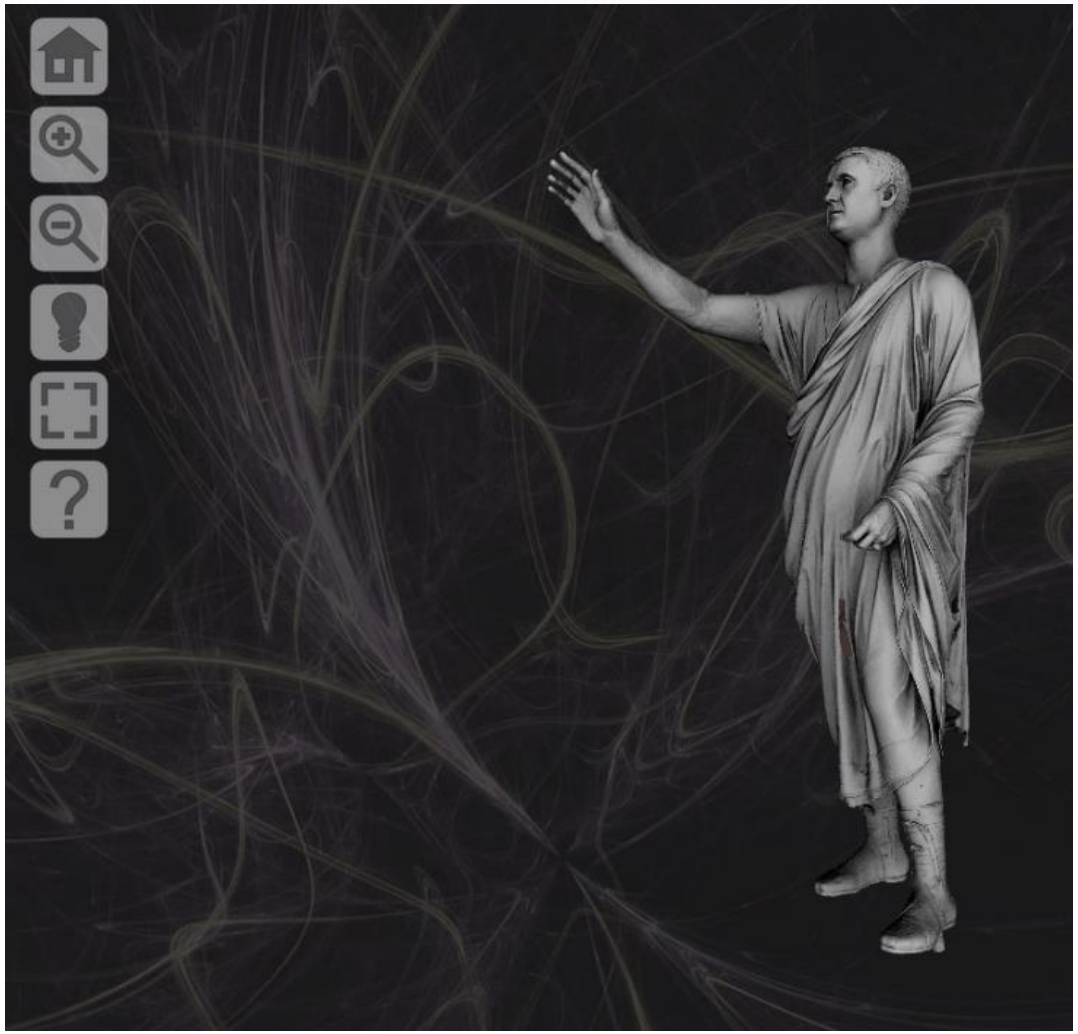


http://visual.ariadne-infrastructure.eu/

Visual media service

Viewing the results 3D:

- 80.23 MB
- 3D model of the Arringatore produced by 3D laser scanning
- National Archaeological Museum and CNR-ISTI



http://visual.ariadne-infrastructure.eu/3d/arringatore

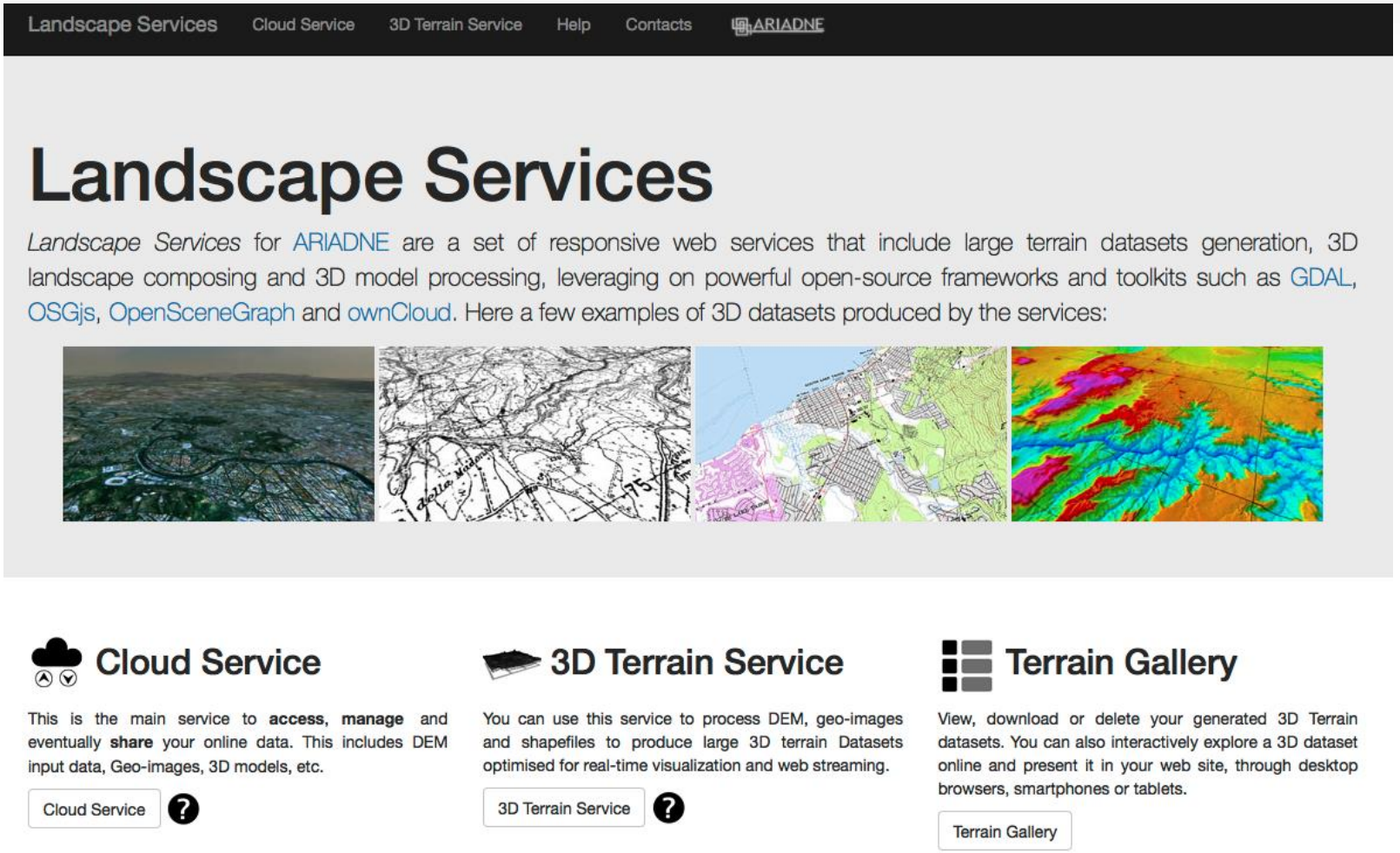
Visual media service

Viewing the results RTI:

- 199.4 mb
- Reverse, Medallion, Rome, Hadrian
- Palazzo Blu, ISTI-CNR, Simoneschi collection

http://visual.ariadne-infrastructure.eu/rti/mediallion_hadrian_reverse

Landscape services

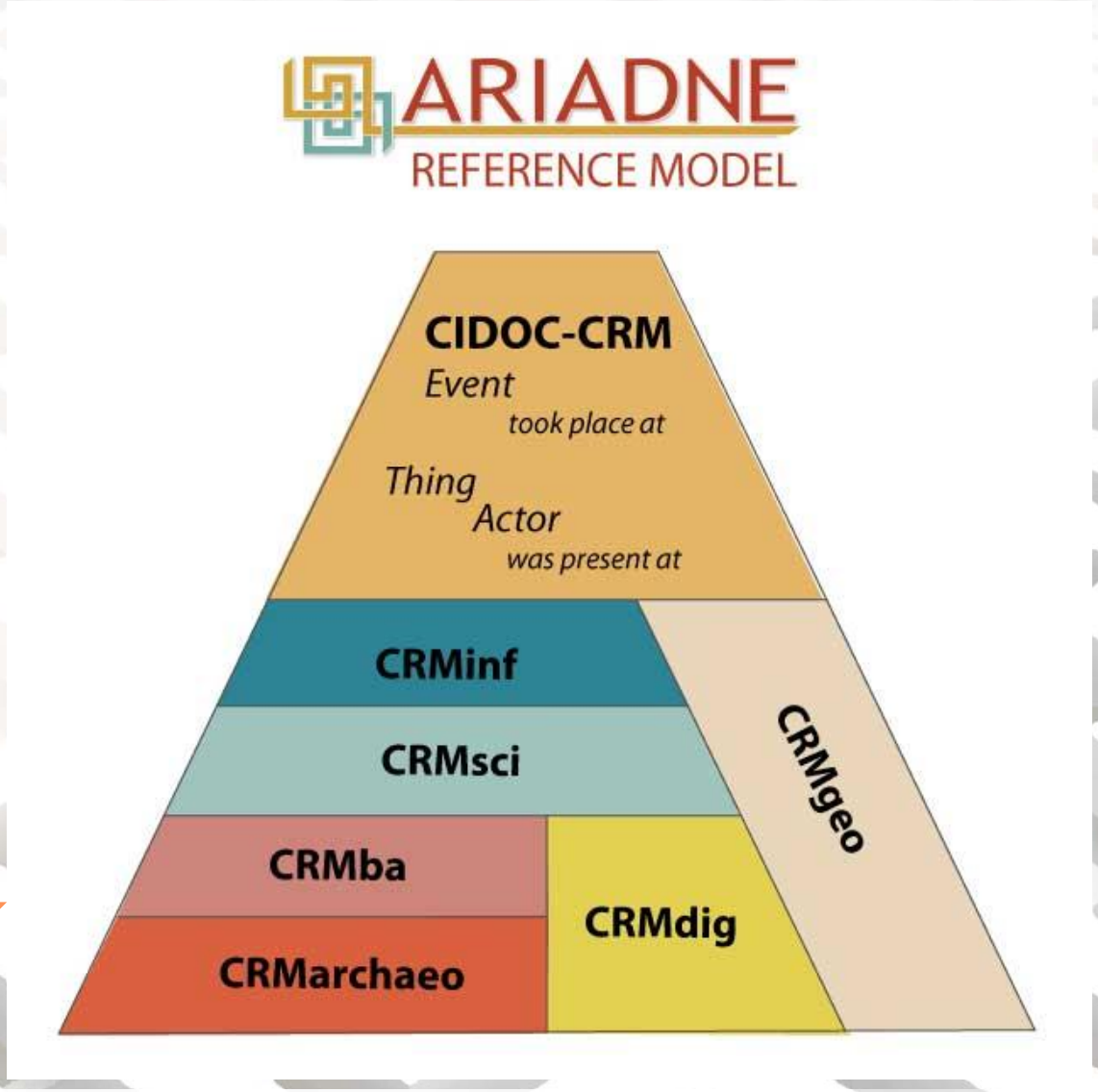


http://landscape.ariadne-infrastructure.eu/

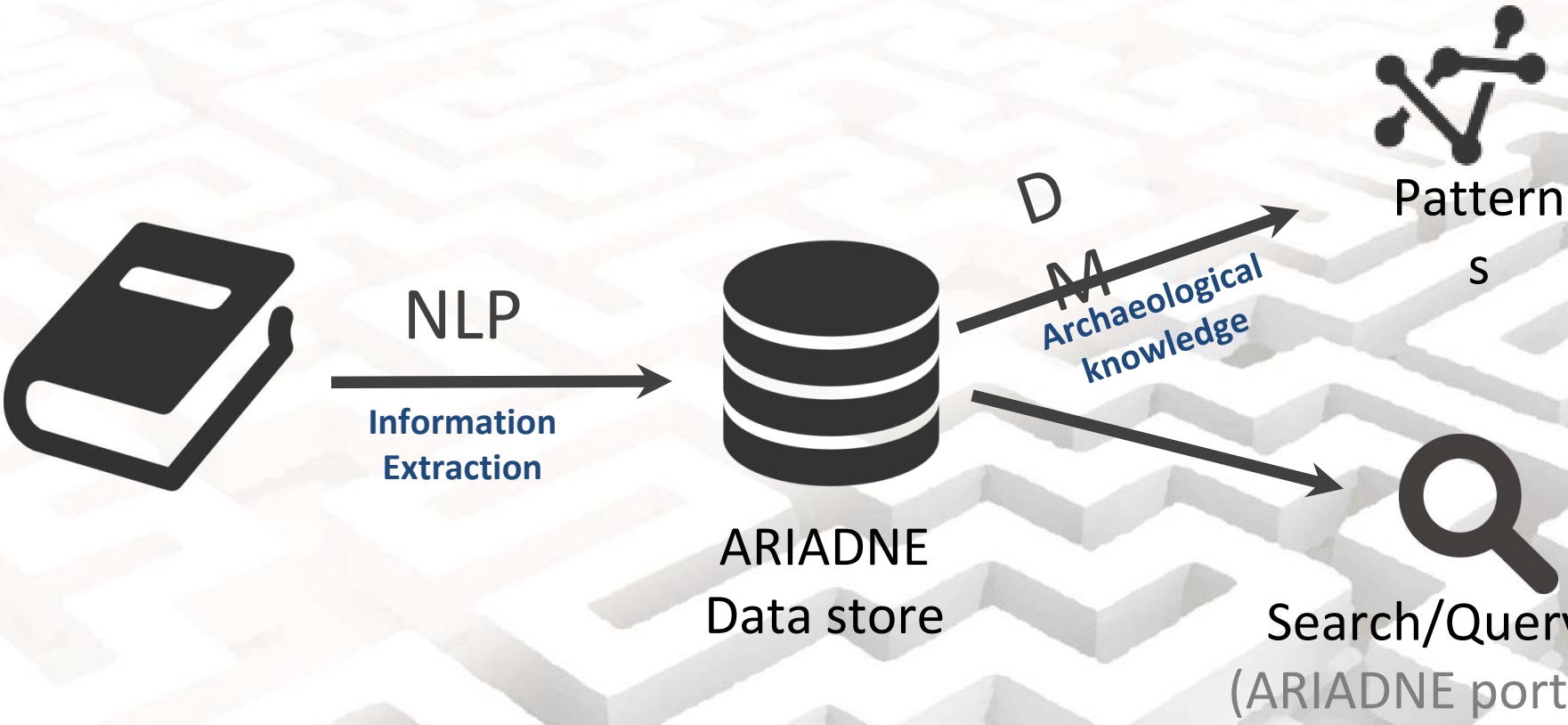
ARIADNE Reference Model

Few concepts, high recall

Special concepts, high precision



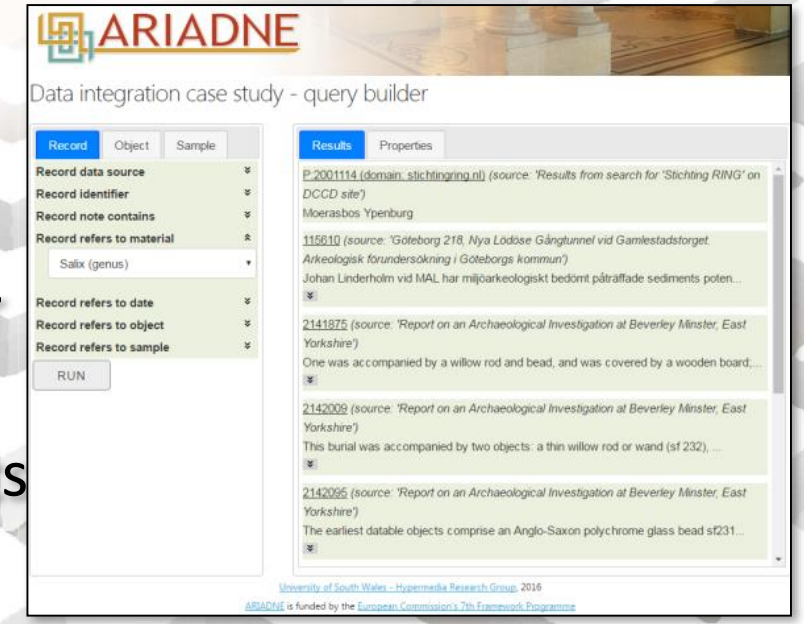
NLP and Data Mining



University of South Wales
Prifysgol De Cymru

Dendrochronology Case Study

- Extracts of 5 archaeological datasets, output from NLP on 25 grey literature reports
- Multilingual - English, Dutch and Swedish data
- Data integration via CIDOC CRM and Getty AAT
- 1.09 million RDF triples
- 23,594 records
- 37,935 objects
- Demonstration query builder for easier cross-search and browse of integrated datasets



Standards: Guides to Good Practice

Archaeology Data Service / Digital Antiquity
Guides to Good Practice

Log in

Dendrochronological Data in Archaeology: A Guide to Good Practice

Peter Brewer, Laboratory of Tree-Ring Research, University of Arizona, USA

Esther Jansma, Cultural Heritage Agency and Utrecht University, The Netherlands

VERSION 1.1 - JUNE 2016

Section 1. Aims and Objectives
- 1.1 Background to the Guide
- 1.2 Scope of the Guide
- 1.3 Data and Metadata

Section 2. Creating Dendrochronological Data
- 2.1 Project Planning and Requirements
- 2.2 Sources of Data
- 2.3 File Types (whilst creating, working with, and processing data)
- 2.4 File Naming Convention
- 2.5 Documenting Data Creation and Processing

Section 3. Archiving Dendrochronological Data
- 3.1 Deciding What to Archive
- 3.2 Deciding How to Archive
- 3.3 Archiving File Types
- 3.4 Converting Data Formats
- 3.5 Archiving Strategies
- 3.6 Metadata and Documentation

Section 4. Copyright
- 4.1 Copyright for Dendrochronology

Data Archiving and Networked Services DANS

ARIADNE

DC 3D Digital Collaboratory for Cultural Dendrochronology

Cultural Heritage Agency Ministry of Education, Culture and Scien

THE UNIVERSITY OF ARIZONA

Universiteit Utrecht

New Guides:

- Dendrochronology
- 3D Models in Archaeology

3D Models in Archaeology: A Guide to Good Practice

Martina Trognitz, IANUS, Deutsches Archäologisches Institut (DAI).
Kieron Niven, Archaeology Data Service.
Valentijn Gilissen, Data Archiving and Networked Services (DANS).

With additional contributions from Ruth Beusing (DAI), Bruno Fanini (CNR), Kate Fernie (2Culture Associates), Roberto Scopigno (CNR), Seta Stuhec (OEAW), and Benjamin Štular (ZRC-SAZU)

2016

Section 1. Aims and Objectives
- 1.1 3D Models in Archaeology
- 1.2 Scope of this Guide
- 1.3 Issues and Concerns

Section 2. Creating 3D Data
- 2.1 Project Planning and Requirements
- 2.2 Sources and Types of 3D Data
- 2.3 File Formats

Section 3. Archiving 3D data
- 3.1 Significant Properties
- 3.2 File types for Archiving and Dissemination
- 3.3 Documentation and Metadata

ARIADNE

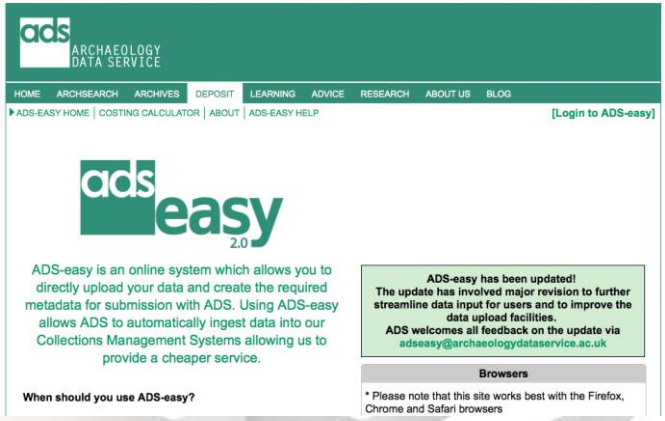
Data Archiving and Networked Services DANS

IANUS

DEUTSCHES ARCHÄOLOGISCHES INSTITUT

Preservation Services

- Care for preservation
- Developing trusted digital repositories:
  - ADS, DANS, IANUS, OAEW
- International collaboration on guidelines



5-10 year Innovation Agenda

- Research e-infrastructures and digital resources for archaeological research and the related domains
- Open sharing and re-use of data: promoting a culture of data sharing, re-use and citation, removing barriers to data sharing
- Data archives for the curation of archaeological research data: reliable and cost-effective community archives for long-term data curation and access
- Capacity building: guidance, training and support for data practices
- Providing services and enabling novel applications

Conclusion

Ariadne was the daughter of Minos, King of Crete.
She gave a ball of thread to Theseus so that he could find his way out of the Minotaur's labyrinth.

Our ARIADNE is a research infrastructure.
We give a portal so researchers can find archaeological datasets and tools to visualise and publish their results.

IMPROVING FINDABILITY AND ACCESSIBILITY

Introduction	Achille Felicetti, PIN
The ARIADNE Common Data Model	Carlo Meghini, CNR
The ARIADNE Registry	Dimitris Gavrilia, ATHENA DCU
Metadata Mappings: What, When & Where	Doug Tudhope, Univ South Wales
The ARIADNE Portal	Sebastian Cuy, DAI
Transnational Activity and training	Achille Felicetti, PIN

IMPROVING INTEROPERABILTY AND REUSABILITY

Introduction

Visual Media Services

Natural Language processing and Data Mining

Linked Data Experiments

Information modeling

Guides to Good Practice

Preservation services

Carlo Meghini, CNR

Roberto Scopingo, CNR

Hans Kamermans, Leiden Univ

Maria Theodoridou, FORTH

César Gonzalez-Perez, CSIC

Holly Wright, ADS, Univ of York

Hella Hollander, DANS

Thanks for your attention

email: info@ariadne-infrastructure.eu

http://www.ariadne-infrastructure.eu
http://portal.ariadne-infrastructure.eu

Twitter: @ariadne-infrastructure

ARIADNE is a project funded by the European Commission under the Community's Seventh Framework Programme, contract no. FP7-INFRASTRUCTURES-2012-1-313193.