

27th November 2013

Partner in charge of the deliverable: **PIN**

Editors for PIN: **Paola Ronzino and Franco Niccolucci**

Quality review: **Julian Richards and Holly Wright, UoY-ADS**

Authors: **Christos Papatheodorou and Dimitris Gavrilis, ATHENA RC**
Kate Fernie, MDR
Holly Wright and Julian Richards, UoY-ADS
Paola Ronzino, PIN
Carlo Meghini, CNR

Contributors: **PIN, UoY-ADS, KNAW-DANS, DAI, MDR, ATHENA RC, CNR, SRFG, Discovery, SND, CSIC, ZRC-SAZU, UoG, MNM-NOK, CYI-STARC, FORTH, ARUP-CAS, OEAW, AIAC, NIAM-BAS, MiBACT-ICCU, ARHEO, INRAP, LU**



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.

Table of Contents

1	Executive Summary.....	6
2	Introduction.....	6
3	Archaeology - oriented metadata standards and conceptual schemas.....	7
3.1	Reference models.....	7
3.1.1	CIDOC CRM	7
3.1.2	CHARM (Cultural Heritage Abstract Reference Model)	8
3.2	Archaeological sites, monuments, landscape areas.....	9
3.2.1	International Core Data Standard for Archaeological Sites and Monuments.....	9
3.2.2	MIDAS Heritage.....	9
3.2.3	CARARE schema	10
3.2.4	ICCD Cataloguing Standards.....	11
3.2.5	EU-CHIC CHICEBERG	13
3.2.6	SIKB0102 - Excavation data.....	13
3.2.7	STARC schema.....	13
3.3	Archaeological sciences.....	13
3.3.1	Tree Ring Data Standard - TRiDaS.....	13
3.4	Museum objects.....	14
3.4.1	SPECTRUM	14
3.4.2	LIDO	14
3.4.3	Object ID	14
3.4.4	VRA Core	15
3.4.5	CDWA.....	15
3.5	Dublin Core.....	15
3.5.1	Dublin Core Application Profiles	16
3.6	Bibliographic materials.....	16
3.6.1	MARC	16
3.6.2	MODS.....	17
3.6.3	METS.....	17
3.7	Archival standards.....	18
3.7.1	EAD	18
3.8	Geospatial information	18
3.8.1	INSPIRE.....	18
3.8.2	CEN/TC 287 Geographic Information.....	18

3.8.3	ISO 19115:2003 - Geographic metadata.....	19
3.8.4	UK GEMINI	19
3.9	Other standards.....	19
3.9.1	Data Documentation Initiative (DDI)	19
3.9.2	Digital Object Identifiers (DOI).....	19
4	Archaeology related terminology resources	20
4.1	International terminology resources.....	20
4.1.1	Art & Architecture Thesaurus (AAT)	20
4.1.2	Cultural Objects Name Authority (CONA).....	20
4.1.3	Union List of Artist Names (ULAN).....	20
4.1.4	DCMI Metadata Terms.....	20
4.1.5	European Language Social Science Thesaurus (ELSST)	20
4.1.6	PACTOLS Thesauri	20
4.2	National terminology resources	21
4.2.1	British Museum Materials Thesaurus - UK.....	21
4.2.2	INSCRIPTION - England	21
4.2.3	RCAHMS Vocabularies - Scotland	23
4.2.4	RCAHMW Vocabularies - Wales.....	23
4.2.5	MiBACT - Italy	23
4.2.6	Referentienetwerk Erfgoed (ABR) - Netherlands.....	24
4.2.7	ZRC SAZU - Slovenia	24
4.2.8	Feldolgo-R - Hungary.....	24
4.2.9	Finnish Ontology Library Service ONKI Vocabularies	24
4.2.10	Museums Vocabularies - Germany	24
4.2.11	Archaeological Dictionary of the DAI - Germany	25
4.3	Geospatial resources	25
4.3.1	Getty Thesaurus of Geographic Names (TGN)	25
4.3.2	Geonames	25
4.3.3	Pleiades Ancient Places Gazetteer.....	25
4.3.4	Pleiades Ancient Places Time Periods Vocabulary	25
4.3.5	Place names database of Ireland	26
5	Metadata registries.....	26
5.1	Definitions	26
5.2	The Metadata Registry Standard ISO/IEC 11179	27

6	Metadata registry standards	29
6.1	The DCAT standard	29
6.2	The ARIADNE Metadata Registry Implementation.....	30
7	The ARIADNE Catalog Data Model.....	33
8	Conclusions.....	34
9	References.....	34

1 Executive Summary

The key to interoperability in any domain is a common set of standards which meet the needs of the sector, and the will to adopt them. This report is the synthesis of an extensive survey of the standards currently developed for use within archaeology, developed by other domains which are relevant to archaeology, and allowing interoperability with other domains useful to archaeology. This report also describes the registry developed within the ARIADNE project [1], which will incorporate the standards determined to be best suited to building the infrastructure. The standards adopted by the project and incorporated into the registry are described in the parallel deliverable within this workpackage: D3.2 Report on project standards. The synthesis of the standards survey is divided into two sections, archaeology-oriented metadata standards and conceptual schemas, and archaeology-related terminology resources. The metadata standards and conceptual schemas consist of the reference models (interlinked sets of defined concepts), the most important of which is the CIDOC-CRM and its extensions. The CIDOC-CRM is the cornerstone ontology for modelling data within the cultural heritage domain, and interoperability with the CRM is an important aspect of many of the other standards. In addition to the reference models, the synthesis includes standards for archaeological sites, monuments and landscapes, archaeological sciences, museum objects, the Dublin Core properties for resource description, bibliographic materials, archival standards and geospatial information.

The terminology resources included in the synthesis are divided into national, international and geospatial categories. While the majority are available in English, irrespective of the country of origin, many have translations into a variety of European languages, which will form a vital basis for interoperability, but translation will not be enough. Terms used within the archaeological domain necessarily vary between languages, as archaeological resources vary from place to place. One of the great challenges to interoperability within archaeology is the difference in the development of human culture in different places at different times. As such, the attempt to make archaeological data from across Europe interoperable represents a significant and unique challenge.

The ARIADNE registry aligned to the standards adopted in D3.2 [2] will begin to address these challenges. The registry described in this report is the ARIADNE Catalogue Data Model (ACDM) and Catalogue system designed and created by ATHENA RC and CNR. The ARIADNE Catalogue is centred on the model of individual datasets, but as many of the partners hold data in the form of collections, the Catalogue has also been extended to handle collections. The Catalogue is also a metadata registry, and this report includes a discussion of metadata registries and standards, along with the technologies used to develop the metadata registry within the Catalogue, followed by a description of the ACDM and Catalogue system itself.

2 Introduction

The main objective of ARIADNE WP3 is to compile a registry of metadata schemas, vocabularies and other standards that are relevant with the Archaeology domain and used by the project's partners, as well as any mappings between them. Due to the large number of standards currently in use, the registry will be an interoperability platform for the research community. The registry will follow the metadata registry standards, such as ISO 11179 and the framework defined by the DESIRE and ROADS projects.

The aim of the first task of the workpackage, Task 3.1, is to survey currently used data standards used by the research community together with any available bilateral mapping. The survey carried out with an extensive analysis of the literature, both on-line and off-line, and directly contacting institutions using those standards if necessary for further information.

The second task, Task 3.2, aims to document the metadata schemas for the ARIADNE's Integrated Infrastructure. Moreover it will also define policies for data access, keeping into account the requirements defined by the owners of rights on the data. Such policies will comply with the EU strategic policies of Open Access to Data.

Finally the third task, Task 3.3, focuses on domain vocabularies and thesauri and will survey and document existing solutions for interoperability, including multilingual aspects. It will also work on gazetteers and other reference lists.

This deliverable is the result of a systematic bibliographic survey of existing metadata standards related to the domain of Archaeology. Significant metadata standards and conceptual schemas are presented from the areas of Archaeology and other relevant fields that are used to describe material related to the archaeological resources, such as bibliographic and archival materials, geospatial information, etc. Moreover, well known existing vocabularies and thesauri are presented. Additionally the report introduces the main concepts related to metadata registries as well as the ISO 11179 standard, as a preliminary background needed for defining the main functionalities of the ARIADNE registry. Finally the ARIADNE Catalog Data Model (ACDM) is briefly introduced as a first draft of the registry, on which the interoperability framework of the ARIADNE project's infrastructure will be developed.

3 Archaeology - oriented metadata standards and conceptual schemas

A working team including UoY-ADS, MDR and PIN completed a survey of standards and schemas used in the archaeology domain. This section describes the major schemas in use.

3.1 Reference models

3.1.1 CIDOC CRM

The CIDOC Conceptual Reference Model (CRM) [3] is a formal ontology whose primary role is to enable information exchange and integration between heterogeneous cultural heritage information. It was developed by the International Committee for Documentation (CIDOC) of the International Council of Museums. In 2006 the International Organization for Standardization adopted it as standard ISO 21127:2006 (Information and Documentation: A Reference Ontology for the Interchange of Cultural Heritage Information). CIDOC CRM aims at providing the semantic definitions and clarifications needed to transform disparate, localised information sources into a coherent global resource, be it within a larger institution, in intranets or on the Internet.

The goal of the CRM is to enable the integration of the largest number of information resources; therefore it aims to provide the greatest flexibility of systems to become compatible, rather than imposing one particular solution.

The CRM does not require complete matching of all user documentation structures with the CRM, nor that systems should always implement all CRM concepts and associations; instead it leaves room both

for extensions, needed to capture the full richness of cultural information, and for simplifications, required for reasons of economy [4].

The CRM is available in English as RDFS and is freely available for download.

CRMdig

CRMdigital [5] is an Extension of CIDOC-CRM to support provenance metadata. It is available in English as RDFS, was created by FORTH, and is freely available for download.

CRMsci

CRMsci: the Scientific Observation Model is an Extension of CIDOC-CRM to support scientific observation. It is available in English as RDFS, was created by FORTH, and will be freely available for download.

CRMgeo

CRMgeo links the CIDOC CRM to GeoSPARQL through a Spatiotemporal Refinement. It is available in English as RDFS, was created by FORTH, and will be freely available for download.

CRM-EH

CRM-EH [6] is an extension to the CIDOC Conceptual Reference Model (CIDOC CRM – ISO 21127:2006 standard), which models the archaeological excavation and analysis process. The work on the CRM-EH was prompted by a need to model the archaeological processes and concepts in use by the English Heritage (EH) archaeological teams, to inform future systems design and to aid in the potential integration of archaeological information in interoperable web based research initiatives.

It is available in English, in the form of an RDF schema created by the University of South Wales in partnership with English Heritage and the Archaeology Data Service, and is freely available for download [7].

3.1.2 CHARM (Cultural Heritage Abstract Reference Model)

CHARM [8] is an abstract and wide model of cultural heritage, aimed at providing a reference conceptual framework. It is designed to be used by a diverse range of people to meet their particular needs. CHARM represents things in terms of concepts, properties of concepts and relationships between concepts. CHARM is a descriptive model and users can pick those elements that meet the needs of their organization and project, creating particular models (extensions of CHARM) in order to use it taking into account the particularities of their situation.

In comparison to the CIDOC CRM, CHARM is wider in scope, encompassing tangible and intangible entities, agents, valorisations, representations, events and other kinds of things related to cultural heritage with equal emphasis. Secondly, CIDOC CRM is a standard while CHARM provides an abstract model that must be extended into particular models to fit each organisation and project. Finally, the language in which CIDOC CRM is expressed has not been explicitly described which makes processing CIDOC CRM models, or extending the standard, difficult. CHARM is expressed in ConML [9], a well-defined conceptual modeling language documented through a public specification.

See CHARM White Paper (<http://www.charminfo.org/Resources/Technical.aspx>) for more information.

CHARM was created by Incipit CSIC with collaborators, and access to the model is freely available.

3.2 Archaeological sites, monuments, landscape areas

3.2.1 International Core Data Standard for Archaeological Sites and Monuments

The Core Data Standard for Archaeological Sites and Monuments [10] is the result of a collaboration between the documentation committee (CIDOC) of the International Council of Museums (ICOM) and the archaeology documentation group of the Council of Europe. The standard has its origins in an international conference of representatives of national archaeological records held in Copenhagen in 1991 where it became clear that there were many similarities in approach for different national records. The decision to develop a core data standard for archaeological sites and monuments was made at the 1992 CIDOC meeting in Quebec, and the Archaeological Sites Working Group was established.

The International Core Data Standard for Archaeological Sites and Monuments of the Architectural Heritage is closely related to the Core Data Index to Historic Buildings and Monuments of the Architectural Heritage so that countries wishing to include all information relating to the man-made environment on one database can do so. The can be linked with standards for movable objects, including the CIDOC standard for archaeological objects CIDOC's International Guidelines for Museum Object Information and Object ID.

The standard specified the minimum categories of information required to make a reasonable assessment of a monument or site, whether for planning, management, academic, or other purposes.

3.2.2 MIDAS Heritage

Midas Heritage [11] is a UK data standard for the historic environment and used for recording information on buildings, archaeological sites, shipwrecks, parks and gardens, battlefields, landscape areas and artefacts. The first edition of the standard was published by the Royal Commission for Historic Monuments for England (RCHME) in 1998. A second edition was published by English Heritage in collaboration with other UK heritage organisations in 2007, and a third edition in 2012. MIDAS is now maintained by English Heritage on behalf of the UK Forum on Information Standards in Heritage (FISH) [12].

MIDAS suggests the minimum level of information needed for recording heritage assets and covers procedures involved in understanding, protecting and managing these assets.

It also covers guidelines on how to support the effective sharing of knowledge, information retrieval and long-term preservation of data.

MIDAS heritage can be described as a set of closely integrated data standards as it is designed to be used in conjunction with SPECTRUM (for artefacts), the UK Gemini Metadata Standard (GIS) [13], the CIDOC CRM (concepts and relationships) and Informing the Future of the Past: Guidelines for Historic Environment Records. It has six main themes each containing several information groups. The six themes are:

- Heritage Asset – Area, Monument, Artefact and Ecofact
- Activities – Investigative Activity, Designation and Protection, Heritage Asset Management activity, Casework and Consultation, Research and Analysis, Historical event
- Information Sources – Archive and Bibliography, Narrative and Synthesis, Management Activity Documentation

- Spatial Information – Location, Map Depiction
- Temporal information – Date and period
- Actor Information – Actor and Role

The standard aims to provide a common information framework and does not cover: what software or file format to use; what to call fields and tables in a database and how they are designed; what indexing terms to use; how to record archives and museum collections, or how to redesign an existing information system.

FISH Interoperability toolkit

The FISH Interoperability Toolkit [14] was developed by the UK Forum on Information Standards in Heritage with funding by English Heritage and the National Trust to assist with the process of moving information between the wide variety of information systems used to record the historic environment.

MIDAS XML is the heart of the Toolkit and is a W3C XML schema, which is freely available for download in English. The toolkit also includes a suite of technical tools, the Historic Environment Exchange Protocol (HEEP) Web Services protocol, documentation and training materials.

3.2.3 CARARE schema

The CARARE schema is a metadata schema based on existing standards from the archaeology and architecture domain. It was established by the CARARE project [15] with version 1.0 being released in autumn 2010 and updated in spring 2011 following testing by content partners. Version 2.0 of the CARARE Schema was released in summer 2013 following work carried out in the 3D-ICONS [16] projects.

The CARARE metadata schema was designed as a harvesting schema intended for use by national heritage agencies and other archaeological organisations to deliver metadata to the CARARE repository in preparation for supply to Europeana [17]. The strength of the schema lies with its ability to support the full range of descriptive information about monuments, building, landscape areas and their 2D and 3D representations and related information sources, collections and events.

The focus of the CARARE schema is on the detailed description of heritage assets (monuments, buildings, landscapes, artifacts and information sources) and related digital resources and events in which the heritage asset is represented. The Schema is based on MIDAS Heritage [18], a data standard for information about the historic environment developed and maintained by English Heritage for the UK Forum on Information Standards in Heritage. The CARARE schema adds elements from LIDO and ESE (Europeana Semantic Elements) to cover the information needed for the digital resources being made accessible online.

Version 2.0 of the CARARE schema includes elements from the CIDOC CRMdig extension to meet the needs of recording provenance and paradata for 3D ICONS objects, and was modified to take into account developments in EDM (Europeana Data Model).

The root element of a CARARE record is the CARARE wrap, which wraps one or many CARARE records. The CARARE schema's core is:

- themes (Heritage Asset, Digital Resource, Activity and Collection information)
- a series of global types (record information, spatial, temporal, rights, appellation, etc) which may be used across the schema to define the information elements

The schema specifies whether themes and elements are mandatory, strongly recommended or optional.

The four major concepts which are wrapped into a main entity – the CARARE record - are:

- Heritage asset – this includes descriptive and administrative metadata for archaeological monuments, historic buildings, industrial monuments, archaeological landscape areas, shipwreck, artifacts and ecofacts; and also images, drawings, plans, maps, archives, publications and 3D models which represent or provide sources of information about a monument, building or landscape. This entity is unique in the CARARE record.
- Digital resource – this provides details about the type, format and online location of a digital file such as a text document, image, audio file, video or 3D object that represents a heritage asset
- Collection information – this describes the collection that holds the content being provided
- Activity – this includes metadata for both historical events, which took place at the heritage assets (such as building, alternation, demolition, battles, etc) and archaeological events (such as excavations, surveys, etc)

The CARARE schema allows for classification by subject and time period, for the identification of place names (spatial coordinates) and agents associated with the site, monument or building, as well as contextual information about activities or events. CARARE entities can be explicitly related to each other. For example there can be relationships between heritage assets (for instance, in the case of a building which is part of a complex), between a heritage asset and the digital resources that can be viewed online, between heritage assets and activities, etc.

CARARE Schema (version 1.0.6)

Version 1.0.6 of the CARARE schema is available in XML in English [19]. Mappings have been completed to Europeana’s ESE and EDM schemas and a partial mapping to the CIDOC CRM is available.

3D-ICONS/CARARE 2.0 Schema

This schema is an update of the CARARE 1.0.6, developed within the 3D-ICONS project, which takes into account the need to capture paradata and provenance data describing the creation of 3D models expressed in the CIDOC CRMdig extension, and developments in EDM. It is available in English as an XML Schema [20]. A mapping for the CARARE 2.0 schema to EDM has been completed, and partial mappings to the CIDOC CRM and CRMdig extension are available.

3.2.4 ICCD Cataloguing Standards

The ICCD (Central Institute for Catalogue and Documentation of the Italian Ministry for cultural heritage, activities and tourism) [21] *corpus* of cataloguing standards consists of regulations, support and control tools (vocabularies, lists of terms) and a set of rules and guidelines illustrating the methods to be followed for the acquisition and production of cultural heritage documentation [22].

The corpus includes:

- *Regulations for cataloguing*, describing the data models and the Authority files to be used for cataloguing activities
- *Catalogue schemas*: descriptive models and forms for collecting information in a structured way, according to a “path of knowledge” [23]
- *Authority files*, to guarantee uniformity in the use of information concerning key concepts (e.g. authors, bibliography) used throughout the whole system. ICCD created and maintains four

Authority files for archaeology, including: “AUT” (Authors), “DSC” (Archaeological Excavations) and “RCG” (Archaeological Surveys)

- *Support and control tools*: thesauri and terminological tools [24]

For archaeology the tools available for cataloguing movable and immovable archaeological properties (according to version 3.00 of the Regulation, recently released) are the following:

SI Schema - Archaeological Sites

Used to describe and document an archaeological site, intended as a “portion of land that preserves evidence of human activities, belonging to a past more or less remote, and investigable with the proper methods of archaeological research”, with any regard to quality, quantity or size of the evidence.

SAS Schema - Stratigraphic Surveys

Used for the documentation of stratigraphic sequences found in contexts of archaeological excavations. The ICCD has an on-going research project for the automatic processing of the records for the detection of Stratigraphic Units.

US Schema – Stratigraphic model

To record the various aspects of archaeological analysis. Under development.

CA Schema - Archaeological Complexes.

Used for the documentation of archaeological properties, regardless of the current state of conservation, having an easily identifiable functional architecture, and composed of various building units (e.g. a fortified place, an insula, etc.).

MA Schema - Archaeological Monuments

Used for the recording of single identifiable building units (a tower, a domus, a temple, etc.), identified and organized on the basis of the functional units (circles) and partitions (walls, roofs, floors, etc.).

TM Schema – Type wall model

To record the various aspects of technical wall. Under development.

RA Schema - Archaeological Finds

Used for the recording of movable objects, it is the most used and well established standard for Italian archaeology, because of the very high number of artefacts as a result of archaeological excavations, surveys and discoveries throughout the national territory.

NU Schema - Numismatics

Used for the recording of objects having a monetary relevance, not only coins but also objects with a monetary connotation, including seals, ancient medals, coinage tools and weights.

TMA Schema - Archaeological Materials

Used for the recording of large collections of materials without significant characteristics or fragmentary, often coming from archaeological excavations or surveys, or stored in museums and private collections, for which it is not expected to use RA schema.

AT Schema - Anthropological Finds

To record biological evidences in close relation with archaeological and paleontological, historical and cultural contexts, affecting the evolution, life and history of studies of the human race and its predecessors.

EP Schema - Epigraphic Model

To record the various aspects of the epigraphic documentation. Under development.

The various standards provide a comprehensive hierarchical framework for top-down analysis (i.e. from the general ‘territorial container’ represented by the archaeological site, through the archaeological complex, the individual monument composed of parts and subparts, to the artefact) and, *vice versa*, to reconstruct the bottom-up sequences from the movable object back to the monumental and territorial context of belonging. There is a strong and articulated system of relationships between the various schemas, which is not rigidly preordained, but can vary to fit different scenarios.

3.2.5 EU-CHIC CHICEBERG

The project EU-CHIC (Cultural Heritage Identity Card) [25] has defined the concept of the CHICEBERG Protocol for the integrated documentation of tangible cultural heritage [26]. The protocol is based on a taxonomy of historic buildings developed within the EU project Perpetuate [27]. The scope of EU-CHIC mainly concerns conservation and documentation of environmental changes affecting tangible cultural heritage assets, buildings and monuments, including "natural" deterioration processes and human interventions.

The data management scheme proposed by EU-CHIC is based on a Core Data Index for architectural heritage and a Core standard for archaeological heritage. It follows the Council of Europe 2009 Guidance on inventory and documentation of cultural heritage [28]. The EU-CHIC scheme is composed by three levels of information, which form the so-called “CHICEBERG”. The first level concerns information describing the asset with free access; the second level concerns the collection of information with restricted or no access where detailed data on the history, architecture, previous interventions, risks and so forth can be stored. The third level of information concerns information about activities that contribute to scientific decision-making support: e.g. preventive conservation regarding monitoring of assets and environmental impact controls, and management through the exploitation and planning of regular and extraordinary maintenance activities.

3.2.6 SIKB0102 - Excavation data

SIKB0102 is the Dutch national standard of exchanging excavation data. It was created as an XML schema, and is available in Dutch. It was created by SIKB, and freely available for download [29].

3.2.7 STARC schema

STARC schema is based on the CARARE (see 2.2.3 above) and LIDO (see 2.3.2 below) schemas and is compliant with the CIDOC-CRM. It was developed under CARARE mainly for the documentation of archaeological assets (objects, ancient buildings, archaeological sites). The goal is to enable data interoperability and access to the digital resources stored in the STARC repository [30]. It is available in English and Greek, was created by Cyl-STARC, and is available as an XML schema [31].

3.3 Archaeological sciences

3.3.1 Tree Ring Data Standard - TRiDaS

TRiDaS is the Tree Ring Data Standard for dendrochronology. It is available in English, was created as part of the TRiDaS project, and is freely available for download as an XML schema [32].

3.4 Museum objects

3.4.1 SPECTRUM

SPECTRUM standard [33] is an open collections management standard, which is widely used for in museums for documentation internationally. Developed and maintained by the Collections Trust (UK) in collaboration with museum professionals the core SPECTRUM standard has been translated in several languages.

3.4.2 LIDO

LIDO is the result of a collaborative effort of international stakeholders in the museum sector, starting in 2008, to create a common solution for organisations to contribute information on their objects to portals or aggregators [34]. LIDO is a harvesting schema designed for delivering metadata from an organisation's online collections databases in a standard way to thematic, cross-domain, national and international portals. The LIDO schema supports the full range of descriptive information about museum objects of all kinds, including art, architecture, cultural history, history of technology, natural history etc. Led by the CDWA Lite Advisory Committee and the Documentation Committee of the German Museums Association, it was agreed to create a single schema that met the requirements articulated by CDWA Lite, museumdat, and feedback received from the greater community of information and technology professionals. A working group was established for the development of LIDO which was joined by representatives from the ATHENA project who helped integrate SPECTRUM requirements into the LIDO schema. Since then it has been used very effectively in many Europeana feeder projects like Athena [35], Linked Heritage [36] and JUDAICA [37].

LIDO only requires 4 mandatory elements (ObjectWorkType, RecordID, RecordSource, Title) and in that way the data providers can decide on how light – or how rich – they want their contributed metadata records to be, while also allowing for delivering data and resources relating to their objects. LIDO can also include links from contributed metadata back to records in the providers' 'home' context while it allows for identification of each referenced entity, e.g. provide references to controlled vocabulary and authority files. In general LIDO structure is divided in the Descriptive and administrative information groups, the first includes metadata about the cultural object while the second one administrative metadata.

LIDO is described as an application of the CIDOC Conceptual Reference Model (CRM) and a mapping between LIDO and the CIDOC CRM is available [38].

A mapping between LIDO and ESE has been completed and work is underway to complete a mapping to EDM [39].

3.4.3 Object ID

Object ID [40] is an international standard for describing cultural objects, launched in 1997 and promoted by major law enforcement agencies, including the FBI, Scotland Yard and Interpol, UNESCO, museums, cultural heritage organizations, art trade and art appraisal organizations, and insurance companies. It provides both a descriptive standard and a means of helping combat the illegal appropriation of art objects.

3.4.4 VRA Core

VRA core [41] is a data standard for the description of works of visual culture and also the images that document them. It consists of a metadata element set (units of information such as title, location, date, etc.), as well as an initial blueprint for how those elements can be hierarchically structured. The element set provides a categorical organization for the description of works of visual culture as well as the images that document them. The latest version of VRA Core 4.0 was produced to make an XML expression possible.

VRA core is hosted by the Network Development and MARC Standards Office of the Library of Congress in partnership with the Visual Resources Association.

3.4.5 CDWA

CDWA [42] (Categories for the Description of Works of Art) provides guidelines for the description of art objects and images. It provides a conceptual framework for art databases for describing and accessing information about works of art, architecture, other material culture, groups and collections of works, and related images.

CDWA includes around 540 categories and subcategories. A small subset of categories are considered core in that they represent the minimum information necessary to identify and describe a work. **CDWA Lite** is an XML schema to describe core records for works of art and material culture based on CDWA and CCO. The CDWA Lite schema was enlarged and integrated into the Lightweight Information Describing Objects (LIDO) schema.

Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images (CCO) includes rules and examples for a core subset of the CDWA categories and the VRA Core Categories.

CDWA has been developed and is maintained by the Getty institute.

3.5 Dublin Core

The **Dublin Core Metadata Element Set** [43] is a set of fifteen properties for use in resource description. It is described as a “core” because the elements are broad and generic; while not specific to the cultural heritage or archaeology sectors, it is widely used within the domain. It was created as an XML schema by DCMI, and is available in English for download [44].

The fifteen core elements are: contributor, coverage, creator, date, description, format, identifier, language, publisher, relation, rights, source, subject, title and type.

These elements are part of a larger set of properties, vocabularies and technical specifications maintained by the Dublin Core Metadata Initiative. **DCMI Metadata Terms** [45] provides an extended set of elements and includes resource classes (including the DCMI Type Vocabulary [DCMI-TYPE]), vocabulary encoding schemes, and syntax encoding schemes.

The elements included in DCMI Terms are: abstract, accessRights, accrualMethod, accrualPeriodicity, accrualPolicy, alternative, audience, available, bibliographicCitation, conformsTo, contributor, coverage, created, creator, date, dateAccepted, dateCopyrighted, dateSubmitted, description, educationLevel, extent, format, hasFormat, hasPart, hasVersion, identifier, instructionalMethod, isFormatOf, isPartOf, isReferencedBy, isReplacedBy, issued, isVersionOf, language, licence, mediator, medium, modified,

provenance, publisher, references, relation, replaces, requires, rights, rightsHolder, source, spatial, subject, tableOfContents, temporal, title, type and valid.

Dublin Core is the result of the Dublin Core Metadata Initiative (DCMI).

3.5.1 Dublin Core Application Profiles

The DCMI has recognized that when it comes to metadata, one size does not fit all and that the metadata needs of particular communities and applications are very diverse. This has resulted in a proliferation of metadata formats including bespoke extensions to the Dublin Core metadata element set.

The DCMI has provided a framework for designing a Dublin Core Application Profile (DCAP) [46]. A DCAP defines metadata records that meet specific application needs while providing semantic interoperability with other applications on the basis of globally defined vocabularies and models.

PICO - Italy

PICO [47] is a cross-domain application profile based on DCMI technical guidelines, conceived for the Portal of Italian Culture, named CulturalItalia.

It includes 18 qualified Dublin Core elements (creator, coverage, description, date, format, identifier, relation, rights holder, subject, type, title) and refinements and encoding schemas (e.g. ISBN, ULAN, etc.).

EASY Metadata - Netherlands

EASY metadata [48] is an adaptation of DCMI Metadata Terms for application-specific schema for DANS internal use. It was created as an XML schema by DANS, is available in English and is freely available for download. An XSLT is available for the mapping of Easy Metadata to CARARE V1.0.6.

EBU Core - Broadcast archives

EBUCore [49] is designed as a minimum and flexible list of attributes to describe audio and video resources for a wide range of broadcasting applications including archives, exchange and production in the context of a Service Oriented Architecture. It is based on the Dublin Core to maximize interoperability with the community of Dublin Core users.

EBUCore expands the list of elements originally defined in EBU Tech 3293-2001 for radio archives, which was also based on Dublin Core. EBUCore 1.3 takes into account latest developments in the Semantic Web and Linked Data community, and is available as a RDF ontology entirely compatible with the W3C Media Annotation Working Group ontology.

EBUCore is developed and maintained by the European Broadcasting Union (EBU) and is mainly used by broadcasters and audio-visual archives.

3.6 Bibliographic materials

3.6.1 MARC

MARC is an acronym for MACHine-Readable Cataloging. It defines a data format that emerged from an initiative led by the US Library of Congress nearly forty years ago. It provides a format standard bibliographic records and related information in a machine-readable form. MARC became USMARC in the 1980s and MARC 21 in the late 1990s. It is maintained by The Network Development and MARC

Standards Office at the Library of Congress and the Standards and the Support Office at the Library and Archives Canada.

MARC21 [50] is supported by the majority of library systems and enables participation in an international bibliographic community following common standards, and the advantage of copy cataloguing at much reduced cost and without need to maintain conversion programs.

MARC XML [51] provides a framework for working with MARC data in an XML environment.

MARC 21 has been mapped to the following metadata standards: MODS, Dublin Core, MARC Character Sets to UCS/Unicode and Digital Geospatial Metadata.

The following metadata standards have been mapped to MARC 21: MODS, Dublin Core, UNIMARC to MARC 21, ONIX and Digital Geospatial Metadata to MARC.

3.6.2 MODS

Metadata Object Description Schema (MODS) [52] is a schema for a bibliographic element set that may be used for a variety of purposes, and particularly for library applications. While not specific to the cultural heritage or archaeology sectors, it is widely used within the domain.

MODS was developed in 2002 by the Library of Congress' Network Development and MARC Standards Office to provide a bibliographic element set that may be used for a variety of purposes, and particularly for library applications. As an XML schema it is intended to be able to carry selected data from existing MARC 21 records as well as to enable the creation of original resource description records.

It includes a subset of MARC fields and uses language-based tags rather than numeric ones, in some cases regrouping elements from the MARC 21 bibliographic format. As of June 2009 this schema is in its third version (version 3.3).

Some advantages of MODS are that the element set is considered to be:

- richer than Dublin Core
- more compatible with library data than ONIX
- more end user oriented than the full MARCXML schema
- simpler than the full MARC format

The standard is maintained by the Network Development and MARC Standards Office of the Library of Congress with input from users. The XML schema is freely available for download in English.

An XSLT for providing a mapping of Easy Metadata to MODS is available.

3.6.3 METS

METS (Metadata Encoding and Transmission Standard schema) [53] is a standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library, expressed using the XML schema language of the World Wide Web Consortium.

The standard is maintained in the Network Development and MARC Standards Office of the Library of Congress, and is being developed as an initiative of the Digital Library Federation.

3.7 Archival standards

3.7.1 EAD

EAD (Encoded Archival Description) [54] is an XML standard for encoding archival finding aids, maintained by the Library of Congress in partnership with the Society of American Archivists. The EAD standard's document type definition (DTD) specifies the elements to be used to describe a manuscript collection as well as the arrangement of those elements (for example, which elements are required, or which are permitted inside which other elements). EAD 1.0 was an SGML DTD; EAD 2002, the second and current incarnation of EAD, was finalized in December 2002 and is an XML DTD. The EAD tag set has 146 elements and is used both to describe a collection as a whole, and also to encode a detailed multi-level inventory of the collection.

Many EAD elements have been, or can be, mapped to content standards (such as DACS and ISAD(G)) and other structural standards (such as MARC or Dublin Core), increasing the flexibility and interoperability of the data.

3.8 Geospatial information

3.8.1 INSPIRE

The INSPIRE directive came into force in May 2007 with the aim of creating a European Union (EU) spatial data infrastructure [55]. The intention is to enable the sharing of environmental spatial information among public sector organisations and to support public access to spatial information. Although not specific to archaeology the INSPIRE directive covers spatial data captured in national inventories of archaeological and architectural sites.

The directive sets out implementation rules for metadata [56], data specifications, network services, data and service sharing, and for reporting. The data specifications include technical guidelines on administrative units, cadastral parcels, geographical names, protected sites, addresses, coordinate reference systems and geographical grid systems [57].

3.8.2 CEN/TC 287 Geographic Information

The European Committee for Standardization, CEN, is multi-sectorial and develops European Standards for most areas. CEN/TC 287 [58] will produce a structured framework of standards and guidelines, which specify a methodology to define, describe and transfer geographic data and services. This work will be carried out in close co-operation with ISO/TC 211.

The standards aim to support the consistent use of geographic information throughout Europe in a way that is compatible with international usage. They will 1) adopt the ISO 19100 series as European standards; 2) develop new standards in cooperation with ISO/TC 211 that are needed for the INSPIRE initiative and other collaborative programmes; 3) facilitate interoperability with related standards initiatives through harmonization and agreements; 4) promote the use of and education on standards on geographic information.

NEN, the Netherlands Institute for Normalisation, is responsible for the secretariat of CEN/TC287.

3.8.3 ISO 19115:2003 - Geographic metadata

ISO 19115:2003 defines the schema required for describing geographic information and services. It provides information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital geographic data [59].

3.8.4 UK GEMINI

A specification for a set of metadata elements for describing geospatial data resources for discovery purposes. It is produced and maintained by the AGI Standards committee. <http://www.agi.org.uk/uk-gemini/>.

3.9 Other standards

3.9.1 Data Documentation Initiative (DDI)

DDI [60] is a metadata specification for the social and behavioural sciences. Expressed in XML, the DDI specification supports the entire research data life cycle including data conceptualization, collection, processing, distribution, discovery, analysis, repurposing, and archiving.

DDI Lifecycle (formerly DDI 13) is designed to document and manage data across its whole lifecycle and provides a modular and extensible framework based on XML schemas. It includes all of the DDI Codebook plus.

DDI Codebook (formerly DDI 12) is a lightweight version of the standard intended for documenting simple survey data.

The DDI Alliance is a self-sustaining membership organization that develops and promotes the DDI specification and associated tools, education, and outreach programs.

3.9.2 Digital Object Identifiers (DOI)

The Digital Object Identifier (DOI®) System [61] provides a framework for persistent identification, managing intellectual content, managing metadata, linking customers with content suppliers, facilitating electronic commerce, and enabling automated management of media. DOI names can be used for any form of management of any data, whether commercial or non-commercial.

DOI names are assigned to any entity for use on digital networks. They are used to provide current information, including where they (or information about them) can be found on the Internet. Information about a digital object may change over time, including where to find it, but its DOI name will not change.

The system is managed by the International DOI Foundation, an open membership consortium including both commercial and non-commercial members, and has recently been accepted for standardisation within ISO.

Over 40 million DOI names have been assigned by DOI System Registration Agencies in the US, Australasia, and Europe. DOI names are widely used in scientific publishing to cite journal articles (more than 98% of all registered DOIs are for scholarly articles).

4 Archaeology related terminology resources

4.1 International terminology resources

4.1.1 Art & Architecture Thesaurus (AAT)

A structured vocabulary used for describing items of art, architecture and material culture. Mainly used by museums, art libraries, archives, cataloguers, and researchers in art and art history [62].

4.1.2 Cultural Objects Name Authority (CONA)

A controlled vocabulary for indexing cultural works such as architecture and movable items, including all types of paintings, sculpture and a wide range of objects [63].

4.1.3 Union List of Artist Names (ULAN)

A controlled list of names of circa 120,000 artists, generally creators of visual arts and architecture, rather than performance artists, actors or dancers [64].

4.1.4 DCMI Metadata Terms

DCMI Metadata Terms [65]. While not specific to the cultural heritage or archaeology sectors, it is widely used within the domain. It was created as an XML schema, and is available in English. It was created by DCMI, and freely available for download.

4.1.5 European Language Social Science Thesaurus (ELSST)

ELSST [66] is a broad-based multilingual thesaurus for the social sciences. Originally based on the monolingual thesaurus Humanities and Social Science Electronic Thesaurus (HASSET) of the UK Data Archive at the University of Essex, ELSST has been developed over the years by the members of the Council of European Social Science Data Archives (CESSDA) for use in an international setting. Its main aim is to facilitate access to data resources across Europe, independent of domain, resource, language and vocabulary. ELSST is used to aid retrieval in the multinational CESSDA data portal.

The source version of ELSST is the English thesaurus. The thesaurus is currently translated into eight other languages: Danish, Finnish, French, German, Greek, Norwegian, Spanish and Swedish. Translation into further languages is planned.

ELSST is a multidisciplinary thesaurus. Coverage is most comprehensive in the core social science disciplines: politics, sociology, economics, education, law, crime, demography, health, employment, and, increasingly, technology.

4.1.6 PACTOLS Thesauri

PACTOLS is comprised of six thesauri, which are polyhierarchical, multilingual and specialised for antiquity and archaeology from prehistory to the Second World War. In addition, there is a mapping between the PACTOLS thesauri and specific vocabularies by INIST-FRANCIS and BAHR, in order to allow

search on DAPHNE [67]. It is available in French, English, German, Spanish, Italian and Dutch, was created by Frantiq [68], and is available in SKOS format [69].

4.2 National terminology resources

4.2.1 British Museum Materials Thesaurus - UK

A hierarchical list of current and historical terms for materials used in museum objects [70] © British Museum.

4.2.2 INSCRIPTION - England

INSCRIPTION [71] is a collection of 'wordlists' maintained or recommended by the Forum on Information Standards in Heritage (FISH). The FISH vocabularies consist of a number of structured terminologies for recording heritage assets such as monuments, buildings, maritime craft, historic aeroplanes, archaeological objects, components and landscapes. They also cover professional management and archaeological event terms (interventions), types of evidence and archaeological sciences (samples, ecofacts). The vocabularies have been developed over many years, in collaboration with organisations such as English Heritage, ALGAO, Collections Trust and the National Trust, and are recommended for use in conjunction with MIDAS Heritage (see 2.2.2 above) [72]. Several of these terminologies have now been made available as Linked Open Data as a result of the SENESCHAL project [73] including:

- Archaeological sciences
- Building materials
- Components
- Event type
- Evidence
- FISH archaeological objects thesaurus
- Maritime craft type
- Monument type
- Period

The full list of INSCRIPTION wordlists can be consulted online: http://www.fish-forum.info/i_lists.htm.

mda Archaeological Objects Vocabulary

This vocabulary was originally developed by the Archaeological Objects Working Party and published by the MDA. It provides guidance for the recording of archaeological objects in Britain and Ireland covering all historical periods. It is now maintained by FISH on behalf of the heritage sector and can be viewed in the English Heritage thesaurus browser [74]. It is also available in SKOS, created by the University of South Wales in partnership with English Heritage and the Archaeology Data Service, and is freely available for download [75].

English Heritage Components Thesaurus

This vocabulary is used to describe the divisions and structural elements of a building or monument and can be consulted online in the EH thesaurus browser [76]. It is also available in the form of a SKOS

vocabulary, created by the University of South Wales in partnership with English Heritage and the Archaeology Data Service, and is freely available for download [77].

English Heritage Building Materials Thesaurus

This vocabulary is used to describe the main constructional material types (e.g. the walls) for indexing of monuments and can be consulted online in the EH thesaurus browser [78]. It is also available in the form of a SKOS vocabulary, created by the University of South Wales in partnership with English Heritage and the Archaeology Data Service, and is freely available for download [79].

English Heritage Event Type Vocabulary

This vocabulary is used to describe the process of recording archaeological and architectural investigative, data collection exercises; from intrusive interventions to non-damaging surveys. It is available in English, in the form of a SKOS vocabulary. It was created by the University of South Wales in partnership with English Heritage and the Archaeology Data Service, and is freely available for download [80].

English Heritage Evidence Vocabulary

This vocabulary is used to describe the existing physical remains of a monument, or the means by which a monument has been identified where no physical remains exist. It can be browsed online in the EH thesaurus browser [81] and is also available in the form of a SKOS vocabulary, created by the University of South Wales in partnership with English Heritage and the Archaeology Data Service, and is freely available for download [82].

English Heritage Monument Type Vocabulary

This vocabulary is a classification of monument type records by function and can be browsed online in the EH thesaurus browser [83]. It is also available in the form of a SKOS vocabulary, was created by the University of South Wales in partnership with English Heritage and the Archaeology Data Service, and is freely available for download [84].

English Heritage Maritime Craft Type Vocabulary

This vocabulary covers inland waterways and sea-going craft types, from warships to simple fishing boats and was developed based on work in compiling English Heritage's maritime record of shipwrecks. It can be browsed online in the EH thesaurus browser [85] and is also available in the form of a SKOS vocabulary, created by the University of South Wales in partnership with English Heritage and the Archaeology Data Service, and is freely available for download [86].

English Heritage Period Vocabulary

This vocabulary describes the periods list used by English Heritage. A version of the list can be browsed on the FISH website [87]. The list is also available in the form of a SKOS vocabulary, created by the University of South Wales in partnership with English Heritage and the Archaeology Data Service, and is freely available for download [88].

English Heritage Archaeological Sciences Thesaurus

This vocabulary is used for recording the techniques, recovery methods and materials associated with archaeological sciences and can be consulted online in the EH Thesaurus Browser [89]. It is also available in the form of a SKOS vocabulary, created by the University of South Wales in partnership with English Heritage and the Archaeology Data Service, and is freely available for download [90].

4.2.3 RCAHMS Vocabularies - Scotland

The Royal Commission on Ancient and Historical Monuments for Scotland (RCAHMS) has developed a series of vocabularies to support its CANMORE database.

RCAHMS Archaeological Objects (Scotland) Vocabulary

This vocabulary describes objects made by human activity as used by the Royal Commission on Ancient and Historical Monuments of Scotland (RCAHMS). It is available in English, in the form of a SKOS vocabulary (created by the University of South Wales in partnership with RCAHMS and the Archaeology Data Service) and is freely available for download [91].

RCAHMS Maritime Craft (Scotland) Vocabulary

This vocabulary describes types of craft that survive as wrecks, or are documented as losses, in Scottish maritime waters as used by the Royal Commission on Ancient and Historical Monuments of Scotland (RCAHMS). It is available in English, in the form of a SKOS vocabulary (created by the University of South Wales in partnership with RCAHMS and the Archaeology Data Service) and is freely available for download [92].

RCAHMS Monument Type (Scotland) Vocabulary

This vocabulary describes monument types relating to the archaeological and built heritage as used by the Royal Commission on Ancient and Historical Monuments of Scotland (RCAHMS). It is available in English with Scottish Gaelic translations for some terms, in the form of a SKOS vocabulary (created by the University of South Wales in partnership with RCAHMS and the Archaeology Data Service) and is freely available for download [93].

4.2.4 RCAHMW Vocabularies - Wales

The Royal Commission on Ancient and Historical Monuments for Wales (RCAHMW) has developed a series of vocabularies to support its Coflein database.

Monument Type (Wales) Vocabulary

This vocabulary describes monument types in Wales by function. It is available in English in the form of a SKOS vocabulary (created by the University of South Wales in partnership with RCAHMS and the Archaeology Data Service) and is freely available for download [94].

RCAHMW Period (Wales) Vocabulary

This vocabulary describes a list of periods for use in Wales. It is available in English in the form of a SKOS vocabulary (created by the University of South Wales in partnership with RCAHMS and the Archaeology Data Service) and is freely available for download [95].

4.2.5 MiBACT - Italy

PICO Thesaurus - Italy

The PICO Thesaurus [96] is a controlled list of hierarchical terms, which covers four main themes: “who”, “what”, “where” and “when”. Its scope includes the cultural heritage, documentary resources, services, domain areas (cultural repositories, artworks, applied art, scientific research, education, events, services and law) and activities for the development and preservation of cultural heritage. It is maintained by MiBACT ICCU for use in the Culturaitalia portal.

Thesaurus per la compilazione delle schede dei Reperti Archeologici

This vocabulary describes a list of terms used for the definition of archeological finds. Most of terms are documented with a relevant picture. It is available in Italian as a PDF. It was created by the Ministero per i beni e le attività culturali, Istituto centrale per il catalogo e la documentazione, and is freely available for download [97].

4.2.6 Referentienetwerk Erfgoed (ABR) - Netherlands

ABR, or the 'Referentienetwerk Erfgoed' (Heritage Network Reference) contains the structured set of concepts of Cultural Heritage in the Netherlands. It is available in Dutch. It was created by the Dutch Cultural Heritage Agency, and can be browsed online via an interface, but an API and SPARQL endpoint is expected to be available in 2014 [98].

4.2.7 ZRC SAZU - Slovenia

ARKAS - Archaeological sites

ARKAS is a list of terms used for the definition of the archaeological sites. Terms are documented by short description. It is available in Slovenian and English, was created by ZRC SAZU, and is available from ZRC SAZU as a text document or for integration into a database.

ZBIVA - archaeological sites, graves, artifacts

ZBIVA is a list of terms used for the definition of the archaeological sites, graves, artifacts. Terms are documented by short description. It is available in Slovenian, German and English, was created by ZRC SAZU, and is available from ZRC SAZU as a text document or for integration into a database.

4.2.8 Feldolg-R - Hungary

Feldolg-R is a database for documenting and cataloguing archaeological finds. Schemes contain: list of types, periods, material, condition and acquisition of objects. It is available in Hungarian, was created by MNM-NÖK, is and XML schema, and is available for integration into a database.

4.2.9 Finnish Ontology Library Service ONKI Vocabularies

The ONKI vocabularies include a general vocabulary, within which there are terms relevant to archaeology, along with a Finnish geography thesaurus. It is available in Finnish, Swedish, and English, was created by Finnish Ontology Library Service (<http://kansalliskirjasto.onki.fi>), and is available in SKOS format [99].

4.2.10 Museums Vocabularies - Germany

This group of vocabularies describes holdings and concepts in use by museums, including a dating system and vessel typology, developed by the Working Group on Documentation within the German Museums Association. It is available in German, in PDF and XML format and is freely available for download [100]. There are additional German language vocabularies relating to museum holdings, which are in SKOS format and based on the CIDOC-CRM, but are not yet published. More information about these vocabularies is available through Jörg Räther at the Archäologisches Museum Hamburg (raether@amh.de).

4.2.11 Archaeological Dictionary of the DAI - Germany

A dictionary currently under development, but a version is available online through a graphical interface, powered by a MySQL database. It is linked to translations in English, French, Italian, Russian, Ukrainian and Chinese, and an English translation links back to the German [101].

4.3 Geospatial resources

4.3.1 Getty Thesaurus of Geographic Names (TGN)

A controlled list of names and associated information about geographical entities, including administrative and political places (e.g. cities, nations) and physical features (e.g., mountains, rivers). Includes current and historical places and information related to history, population, culture, art and architecture [102].

4.3.2 Geonames

The GeoNames [103] geographical database covers all countries and contains over eight million place names that are available for download free of charge under a creative commons attribution licence. It contains over 10 million geographical names and consists of over 8 million unique features whereof 2.8 million populated places and 5.5 million alternate names.

4.3.3 Pleiades Ancient Places Gazetteer

Pleiades provides historical geographic information [104] about the ancient world, and has coverage for the Greek and Roman world, and expanding into Ancient Near Eastern, Byzantine, Celtic, and Early Medieval geography. 'Pleiades concepts are somewhat different from those of other conceptual systems in the cultural heritage and geographic information domains. The entity of the CIDOC Conceptual Reference Model (CRM) labeled E53 Place...is almost exactly equivalent to the Pleiades concept of Location. The Pleiades Place has no single equivalent entity in the CRM. Many places are localized (settlements, stations, temples and monuments) and have much in common with the CRM's E27 Site. Others like ethnic territories, areas of centuriation, or mining districts are rather different. The concept of Feature in the OGC/ISO 19101 system on which many enterprise GIS systems are founded overlaps with both Pleiades Location and to a lesser degree Place. Pleiades Names are nothing like ISO 19101 Features and, as entities in their own rights, are much more than simple feature attributes' (<http://pleiades.stoa.org>). Pleiades was created by the Ancient World Mapping Center and Institute for the Study of the Ancient World, includes a wide variety of place names in both ancient and modern languages, and is freely available for download in CSV, KML and RDF format.

4.3.4 Pleiades Ancient Places Time Periods Vocabulary

This vocabulary consists of a list of archaeological time periods correlated with their associated ancient place [105]. The list is available to view.

4.3.5 Place names database of Ireland

The database is developed by Fiontar (DCU) and The Placenames Branch (Department of Arts, Heritage and the Gaeltacht). It contains details of 61,000 town lands plus details of streets, districts, parishes and other geographic units and can be browsed online [106].

5 Metadata registries

5.1 Definitions

According to Wikipedia “A metadata registry is a central location in an organization where metadata definitions are stored and maintained in a controlled method”.

Data must be portable and shareable within and among various application environments. Therefore, both the users and owners of data must have a common understanding of its meaning, representation, and identification. To understand the meaning of any data, the descriptions of the data must be available from Data Element Registries (Metadata Registries). Data must be adequately described and users must have a convenient way to obtain these descriptions. Data Element Registries provide a way to organize the content and representation of data elements so that data descriptions are consistently specified and can be easily located by data designers and users. Uniform specification of data facilitates data retrieval, data exchange, and consistent use of data throughout the Software Development Life Cycle. The units of information with normalized meanings and formats are known as "standardized data elements." [107].

A Metadata Registry is an information source maintained by a Registration Agency, which describes the structure, the data types, the representation and the semantics of the content of metadata elements. In fact it is a catalogue that contains information (metadata) for metadata elements. Indicatively, the information that a registry holds for a metadata element are: identifier, name, definition and scope note, allowed values, thematic categories, context and application environment and administrative information, such as the submission agent (the agent/organization who registered the element) dates, status of the description, etc.

In particular each element is described by a set of classes of attributes. The main classes of attributes are:

- Identification attributes: name, id, version, equivalent elements, registration agency, context, etc.
- Definition
- Relations: Each element is classified thematically and is correlated to a concept of a classification system. The main attributes of this set are: type of the relation (e.g. ‘qualifier of’, ‘qualified by’, ‘subject of’, ‘part of’, ‘external reference’, ‘higher standard’, ‘data element concept’), the name of the classification system, keywords, references to other metadata elements, etc.
- Representation attributes: Each element has a datatype and a representation format. The members of this attribute set are: representation category (e.g. graphic, barcode, identifier, number, etc.), datatype, min and max values, allowed values, etc.
- Administrative attributes: submission agency, responsible agency, status of the element description, etc.

In conclusion, a Metadata Registry defines a framework in which metadata elements are identified and described uniformly and thus they can be discovered, retrieved, re-used and managed by several different applications and services related to a domain that need to interoperate. Hence the registry provides a knowledge organization system to organize the metadata elements.

This organization enables the discovery of thematically relevant elements, so that the development of application profiles to be facilitated by re-using elements, especially in domains where there exist several metadata schemas and the adoption of common standards is not a frequent practice.

Some well-known metadata registries are:

- Agency for Healthcare Research and Quality- United States Health Information Knowledgebase (USHIK)
- Dublin Core Metadata Registry
- Cancer Data Standards Repository
- National Information Exchange Model
- NIST ebXML Registry for HL7 / HIMSS / IHE
- Open Metadata Registry (formerly the National Science Digital Library (NSDL) Metadata Registry)
- US Environmental Protection Agency - Environmental Data Registry

5.2 The Metadata Registry Standard ISO/IEC 11179

ISO/IEC 11179 describes the standardizing and registering of data elements to make data understandable and shareable. Data element standardization and registration as described in ISO/IEC 11179 allows the creation of a shared data environment in much less time and with much less effort than it takes for conventional data management methodologies.

The purpose of ISO/IEC 11179 is to give concrete guidance on the formulation and maintenance of discrete data element descriptions and semantic content (metadata) that shall be used to formulate data elements in a consistent, standard manner. It also provides guidance for establishing a data element registry. Although motivated by the desire for the open exchange of data throughout the international communities by electronic information interchanges, ISO/IEC 11179 [108]:

- facilitates acquisition and registration of data
- expedites access and use of data
- simplifies data manipulation by intelligent software by enabling manipulation of data based on characteristics described by metadata
- enables the development of a data representation metamodel for CASE tools and repositories
- facilitates electronic data interchange and data sharing

ISO/IEC 11179 benefits the communication of data among information systems and people:

within an organization; among different organizations; and crossing all levels of software and hardware, and geographic, organizational and political boundaries.

Metadata about data elements is stored in a data element registry. Registration is the process of documenting metadata to support data shareability. Registration should be carried out at the data

element level to promote and maximize semantic value. ISO/IEC 11179 enables the end user to interpret the intended meaning confidently, correctly, and unambiguously.

ISO/IEC 11179 - Specification and Standardization of Data Elements - is divided into six parts. The names of the parts, a short description of each, and the status follow below:

Part 1 - *Framework for the Specification and Standardization of Data Elements*: Provides an overview data elements and the concepts used in the rest of the standard.

Part 2 - *Classification of Data Elements*: Describes how to classify data elements.

Part 3 - *Basic Attributes of Data Elements*: Defines the basic set of metadata for describing a data element.

Part 4 - *Rules and Guidelines for the Formulation of Data Definitions*: Specifies rules and guidelines for building definitions of data elements.

Part 5 - *Naming and Identification Principles for Data Elements*: Specifies rules and guidelines for naming and designing non-intelligent identifiers for data elements.

Part 6 - *Registration of Data Elements*: Describes the functions and rules that govern a data element registration authority.

ISO/IEC 11179-3 (Part 3) aims to define a conceptual model for a metadata registry in order to provide a framework for how data elements are formed and the relationships among the parts. The main concepts of the ISO/IEC 11179-3 (Part 3) conceptual model are [109]:

- attribute: A characteristic of an object or entity
- attribute value: A representation of an instance of an attribute
- data element: A unit of data for which the definition, identification, representation and permissible values are specified by means of a set of attributes
- data element concept: A concept which can be represented in the form of a data element, described independently of any particular representation
- data element dictionary: An information resource that specifies, defines, and lists all relevant data elements
- data element value: A value out of a set of permissible values pertaining to a data element

Another important feature of the standard concerns the registration of metadata. Registration is a process that administers metadata; it keeps track of who submitted the metadata, who is responsible for it, and the quality of the metadata provided. Moreover Part 3 contains a common metadata entity called `administered_` component, which captures the metadata common to all objects described in the registry. From the perspective of the common metadata, the `administered_` component entity is like a library card catalog. This makes searching for some objects much easier.

6 Metadata registry standards

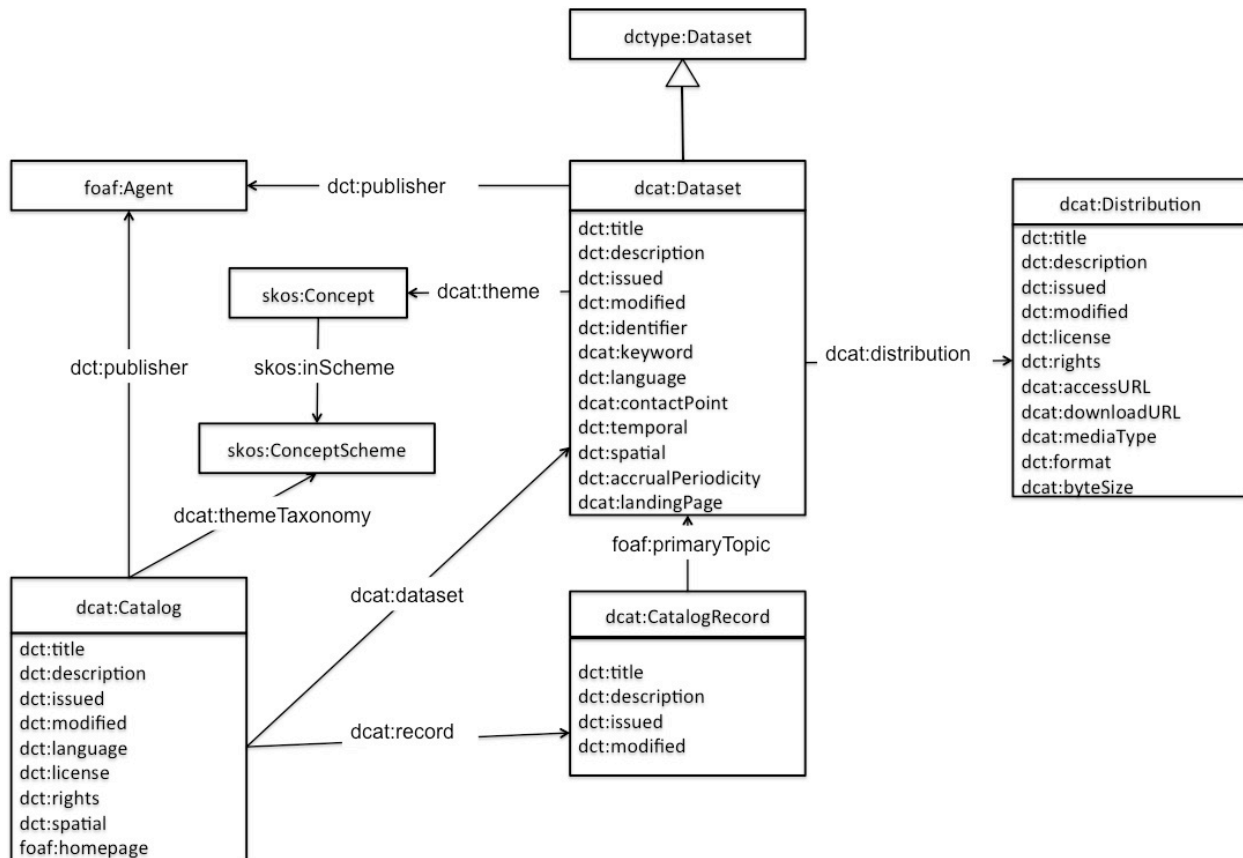
6.1 The DCAT standard

The metadata registry has been based on the Data Catalog Vocabulary standard (DCAT), an RDF vocabulary designed to facilitate interoperability among data catalogs. The DCAT model is used to describe datasets that are found in data catalogs in order to enable their discoverability and consumption by services. The DCAT standard uses the following namespaces:

Prefix	Namespace
dcat	http://www.w3.org/ns/dcat#
dct	http://purl.org/dc/terms/
dctype	http://purl.org/dc/dcmitype/
foaf	http://xmlns.com/foaf/0.1/
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
rdfs	http://www.w3.org/2000/01/rdf-schema#
skos	http://www.w3.org/2004/02/skos/core#
vcard	http://www.w3.org/2006/vcard/ns#
xsd	http://www.w3.org/2001/XMLSchema#

It assumes that a data catalogue conforms to the schema only if it is organized into datasets and distributions.

DCAT brings a number of classes such as foaf:Agents, Concepts, Distributions, CatalogRecords, a set of properties for each and a set of relations among them. The following figure gives an overview of the main interactions among the basic classes of DCAT. More information on DCAT can be found at: <http://www.w3.org/TR/vocab-dcat/>



6.2 The ARIADNE Metadata Registry Implementation

The metadata registry has been implemented using web based technologies and more specifically using PHP 5.x for the frontend. It uses a MySQL 5.x database and also uses a set of Java SE applications in order to perform certain functionalities. The development pattern adopted using the implementation of the metadata registry was based on the following requirements:

- It represented each class of the ACDM model separately and allowed the easy management of each class properties
- It allowed the interlinking of different class instances to be modeled as properties.

Development cycle

This model resembled an agile programming like development as it allowed the implementation to produce a working prototype quickly and also allow it to follow the continuous changes of the ACDM model. Each property has a specific data type that can be one of:

- simple input field
- text field (textarea with multiple rows)
- select list
- file attachment
- autocomplete field

- link to another class instance

Elements can be defined as repeatable.

Interfaces

The metadata registry provides three (3) different interfaces for managing information:

1. Web based UI. This is obviously designed for manual editing of information by human users. Each content provider has his own user account through which he can add new / edit existing information per class and link different instances together.
2. Batch Import. The batch import interface is controlled through a web based ingest control panel and allows the batch ingest of datasets and collections through: a) OAI-PMH target, b) XML file, c) Excel file. The ingest tool comes with an embedded mapping tool and allows to map any xml/excel structure to a dataset or collection.
3. REST API. The REST API allows to export all catalog information (per provider) in XML format and also to manipulate it through a REST API that is provided. The access to the API is controlled by a session key that is generated from within the Web interface (per provider).

Linked data

The registry follows the linked data guidelines and exports in XML or RDF using the provided REST API. The namespace used for the extension defined within the ACDM document is: ***ariadne:*** whereas the identifiers for each instance start with a prefix representing the class they belong to (e.g. dis: for distribution and agn: for agent). In the future (Workpackages 12 and 13) this will allow ARIADNE to provide a more robust and public REST interface (possibly a SPARQL endpoint). These screenshots of the system implementation illustrate the design concept:

About Ariadne



ARIADNE brings together and integrates existing archaeological research data infrastructures so that researchers can use the various distributed datasets and new and powerful technologies as an integral component of the archaeological research methodology. There is now a large availability of archaeological digital datasets that all together span different periods, domains and regions; more are continuously created as a result of the increasing use of IT. The are the accumulated outcome of the research of individuals, teams and institutions, but form a vast and fragmented corpus and their potential has been constrained by difficult access and non-homogenous perspectives.

[Find out more](#)

Ariadne Metadata Registry

The ARIADNE metadata registry will survey currently used data standards and metadata schemas, and will compile a registry of those in use in different fields/regions of archaeological research, including important related domains of research, together with any available bilateral mapping.

[Login](#)

Screenshot showing the home page of the metadata editor

← ARIADNEDataset

dct:title	<input type="text"/>	dct:publisher	<input type="text"/>
dct:description	<input type="text"/>	dct:creator	<input type="text"/>
dct:issued	<input type="text"/>	:owner	<input type="text"/>
dct:modified	<input type="text"/>	:responsible	<input type="text"/>
:originalId	<input type="text"/>	dcat:theme	<input type="text"/>
dct:identifier	<input type="text"/>	:hasAttDocuments	<input type="text"/>
dcat:keyword	<input type="text"/>		
dct:language	<input type="text" value="en"/>		

Screenshot showing a part of the Dataset cataloguing form

7 The ARIADNE Catalog Data Model

The ARIADNE registry conceptual model is named ACDM. It is an extension of the Data Catalog Vocabulary (DCAT), a quasi-recommendation of the W3C Consortium [110] that *“is well-suited to representing government data catalogs such as Data.gov and data.gov.uk.”* The reason for adopting the DCAT Vocabulary (apart from re-use) is that DCAT is proposed as a tool for publishing datasets as Open Data. Its adoption places therefore ARIADNE in an ideal position for publishing datasets as Open Data as well. To this end, ARIADNE will be following the recommendations made in the “DCAT Application Profile for data portals in Europe” concerning the use of the terms of the DCAT ontology. These recommendations establish which attributes or classes are mandatory; for the moment, we will not strictly adhere to these recommendations because we are using DCAT for internal purposes.

The central notion of the model is the class `ArchaeologicalResource` that has as instances the main resources described in a catalogue. These are categorized in:

- services, instances of the class `Service`
- language resources, instances of the class `LanguageResources`, further categorized in vocabularies, mappings (between metadata formats) and gazetteers. This is just a preliminary categorization to be further expanded as new resources of linguistic natures are introduced into the catalogue (such as subject heading systems, ontologies and the like)
- data resources, instances of the class `DataResources`, further categorized in collections and datasets, the latter categorized in datasets proper, databases and GIS
- metadata schemas, instances of the class `MetadataSchema`

A very important class of the model is that of formats, called `MetadataFormat`. Instances of this class are the formats of various resources described in the catalogue, namely:

- collection- or dataset-level metadata,
- archaeological records (i.e., members of datasets)
- metadata of archaeological records
- metadata of collection items.

Formats are represented in the catalogue in order to provide information useful for integration.

Finally, there is a class for representing collection- or dataset-level metadata records, which are stored in the catalogue in order to support discovery of similar resources that are natural candidates for integration.

There are several associations that link these classes between each other and to other, less important classes, such as agents or concepts. The model is attached as an appendix of this deliverable.

8 Conclusions

This report has provided an extensive survey of the standards currently developed for use within archaeology developed by other domains, which are relevant to archaeology, and allowing interoperability with other domains useful to archaeology. These include archaeology-oriented metadata standards and conceptual schemas, and archaeology-related terminology resources. The metadata standards and conceptual schemas consist of the reference models (interlinked sets of defined concepts), the most important of which is the CIDOC-CRM and its extensions. The CIDOC-CRM provides the cornerstone ontology for modelling data within the cultural heritage domain, and interoperability with the CRM will be an important requirement for the ARIADNE data model. The survey has also revealed a broad range of standards for archaeological sites, monuments and landscapes, archaeological sciences, museum objects, the Dublin Core properties for resource description, bibliographic materials, archival standards and geospatial information. While the majority are available in English, irrespective of the country of origin, many have translations into a variety of European languages, which will form a vital basis for interoperability, but mere translation will not be enough. Terms used within the archaeological domain necessarily vary between languages, as archaeological resources vary from place to place. One of the great challenges to interoperability within archaeology is the difference in the development of human culture in different places at different times. Virtually all parts of Europe have a period they would refer to as being ‘Neolithic’ but when that period is generally agreed to have occurred is hugely varied depending on what part of Europe is under study. Equally, the regions of Europe, which were not part of the Roman Empire will have no Roman artefacts. As such, the attempt to make archaeological data from across Europe interoperable represents a significant and unique challenge. The ARIADNE registry aligned to the standards adopted in D3.2 will begin to address these challenges. The registry described in this report is the ARIADNE Catalogue Data Model (ACDM) and Catalogue system designed and created by ATHENA RC and CNR. The ARIADNE Catalogue is centred on the model of individual datasets, but as many of the partners hold data in the form of collections, the Catalogue has also been extended to handle collections.

9 References

1. ARIADNE project <http://www.ariadne-infrastructure.eu/>
2. D3.2 Report on project standards, on ARIADNE web site www.ariadne-infrastructure.eu/
3. CIDOC CRM <http://www.cidoc-crm.org/cidoc-crm> (consulted 28/10/13)
4. Nick Crofts, Martin Doerr, Tony Gill, Stephen Stead, Matthew Stiff (editors), Definition of the CIDOC Conceptual Reference Model, December 2011.
5. http://www.ics.forth.gr/isl/CRMext/CRMdig_v3.0.rdfs (consulted 28/10/13)
6. <http://hypermedia.research.southwales.ac.uk/resources/crm/> (consulted 21/10/13)
7. http://www.heritagedata.org/crmeh/crmeh_current.rdf (consulted 28/10/13)
8. <http://www.charm.info.org/Reference/> (consulted 21/10/13)
9. <http://www.conml.org/> (consulted 21/10/13)
10. <http://archives.icom.museum/objectid/heritage/intro4.html>
11. <http://www.english-heritage.org.uk/publications/midas-heritage/>

12. <http://fishforum.weebly.com/>
13. UK Gemini: <http://www.agi.org.uk/uk-gemini/>
14. <http://www.heritage-standards.org.uk/>
15. CARARE project website, <http://www.carare.eu>
16. 3D ICONS project website, <http://www.3DICONs-project.eu>
17. Europeana, <http://www.europeana.eu>
18. MIDAS Heritage, <http://www.english-heritage.org.uk/publications/midas-heritage/>
19. <http://www.carare.eu/eng/Resources/CARARE-Documentation/CARARE-metadata-schema> (consulted 1/11/13)
20. <http://www.carare.eu/eng/content/download/11455/98743/file/carare-v2.0.1.xsd> (consulted 1/11/13)
21. <http://www.iccd.beniculturali.it/>
22. A. Felicetti, T. Scarselli, M. L. Mancinelli, F. Niccolucci, 2013. Mapping ICCD Archaeological Data to CIDOC-CRM: the RA Schema (in press)
23. <http://www.iccd.beniculturali.it/index.php?it/204/normative>
24. <http://www.iccd.beniculturali.it/index.php?it/240/vocabolari>
25. www.eu-chic.eu
26. <http://eu-chic.eu/index.php/news/entry/chiceberg/>
27. www.perpetuate.eu
28. Council of Europe 2009 *Guidance on inventory and documentation of the cultural heritage*. Published by the CoE.
29. http://www.sikb.nl/upload/documents/BRL0100/SIKB0102-2.1.0_XSD.zip (consulted 1/11/13)
30. <http://public.cyi.ac.cy/starcRepo/> (consulted 1/11/13)
31. http://www.academia.edu/1227094/A_Metadata_Schema_for_Cultural_Heritage_Documentation
32. <http://www.tridas.org/download.php> (consulted 1/11/13)
33. <http://www.collectionslink.org.uk/programmes/spectrum>
34. <http://network.icom.museum/cidoc/working-groups/data-harvesting-and-interchange/what-is-lido/>
35. ATHENA project website: <http://www.athenaeurope.org/>
36. Linked Heritage project website: <http://www.linkedheritage.eu/>
37. JUDAICA project website: <http://www.judaica-europeana.eu/>
38. http://www.cidoc-crm.org/crm_mappings.html
39. <http://www.athenaeurope.org/index.php?en/1/home>
40. <http://archives.icom.museum/objectid/about.html>
41. <http://www.loc.gov/standards/vracore/>
42. http://www.getty.edu/research/publications/electronic_publications/cdwa/
43. <http://dublincore.org/documents/dces/> (consulted 21/10/13)
44. <http://dublincore.org/schemas/xmls/qdc/dc.xsd> (consulted 1/11/13)
45. <http://dublincore.org/documents/dcmi-terms/>
46. <http://dublincore.org/documents/profile-guidelines/> (consulted 21/10/13)
47. <http://www.culturaitalia.it/opencms/export/sites/culturaitalia/attachments/documenti/picoap/picoap1.0.xml> (consulted 21/10/13)
48. <http://easy.dans.knaw.nl/schemas/md/emd/2012/11/emd.xsd> (consulted 1/11/13)
49. <http://tech.ebu.ch/lang/en/MetadataEbuCore> (consulted 21/10/13)
50. <http://www.loc.gov/marc/bibliographic/> (consulted 21/10/13)
51. <http://www.loc.gov/marc/marcxml.html> (consulted 21/10/13)

52. <http://www.loc.gov/standards/mods/mods-overview.html> (consulted 1/11/13)
53. <http://www.loc.gov/standards/mets/>
54. <http://www.loc.gov/ead/>
55. INSPIRE directive: <http://inspire.jrc.ec.europa.eu/index.cfm/pageid/48> (consulted 5/10/2013)
56. INSPIRE metadata: <http://inspire.jrc.ec.europa.eu/index.cfm/pageid/101> (consulted 5/10/2013)
57. INSPIRE data specifications: <http://inspire.jrc.ec.europa.eu/index.cfm/pageid/2> (consulted 5/10/2013)
58. <http://www.centc287.org/> (consulted 5/10/2013)
59. http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020
60. <http://www.ddialliance.org/what> (consulted 24/11/2013)
61. <http://www.datacite.org/whatisdoi>
62. © Getty Research Institute/ J.Paul Getty Trust
<http://www.getty.edu/research/tools/vocabularies/aat/index.html> (consulted 21/10/13)
63. © Getty Research Institute/ J.Paul Getty Trust
<http://www.getty.edu/research/tools/vocabularies/cona/index.html> (consulted 21/10/13)
64. © Getty Research Institute/ J.Paul Getty Trust
<http://www.getty.edu/research/tools/vocabularies/ulan/index.html> (consulted 21/10/13)
65. <http://dublincore.org/schemas/xmls/qdc/dcterms.xsd> (consulted 1/11/13)
66. <http://elsst.esds.ac.uk/Home.aspx>
67. <http://www.daphne.cnrs.fr>
68. <http://www.frantiq.fr>
69. <http://pactols.frantiq.fr/opentheso/> (consulted 1/11/13)
70. <http://www.collectionslink.org.uk/discover/manage-information/1136-british-museum-materials-thesaurus> (consulted 21/10/13)
71. <http://www.fish-forum.info/inscript.htm> (consulted 22/11/2013)
72. <http://fishforum.weebly.com/fish-terminology.html> (consulted 22/11/2013)
73. <http://fishforum.weebly.com/linked-open-data.html> and <http://www.heritagedata.org> (consulted 22/11/2013)
74. mda archaeological objects thesaurus: http://thesaurus.english-heritage.org.uk/thesaurus.asp?thes_no=144&thes_name=MDA%20Object%20Type%20Thesaurus
75. http://heritagedata.org/test/seneschal_mda_obj_20130717.rdf (consulted 28/10/13)
76. EH Components: http://thesaurus.english-heritage.org.uk/thesaurus.asp?thes_no=546
77. http://heritagedata.org/test/seneschal_ah_com_20130717.rdf (consulted 28/10/13)
78. http://thesaurus.english-heritage.org.uk/thesaurus.asp?thes_no=129
79. http://heritagedata.org/test/seneschal_ah_tbm_20130717.rdf (consulted 28/10/13)
80. http://heritagedata.org/test/seneschal_agl_et_20130717.rdf (consulted 28/10/13)
81. http://thesaurus.english-heritage.org.uk/thesaurus.asp?thes_no=92
82. http://heritagedata.org/test/seneschal_ah_evd_20130717.rdf (consulted 28/10/13)
83. http://thesaurus.english-heritage.org.uk/thesaurus.asp?thes_no=1
84. http://heritagedata.org/test/seneschal_ah_tmt2_20130717.rdf (consulted 28/10/13)
85. http://thesaurus.english-heritage.org.uk/thesaurus.asp?thes_no=143
86. http://heritagedata.org/test/seneschal_ah_tmc_20130717.rdf (consulted 28/10/13)
87. http://www.fish-forum.info/i_apl_e.htm
88. http://heritagedata.org/test/seneschal_ah_period_20130717.rdf (consulted 28/10/13)
89. EH Archaeological Sciences: http://thesaurus.english-heritage.org.uk/thesaurus.asp?thes_no=560
90. http://heritagedata.org/test/seneschal_ah_560_20130717.rdf (consulted 28/10/13)

91. http://heritagedata.org/test/seneschal_rc_2_20130717.rdf (consulted 28/10/13)
92. http://heritagedata.org/test/seneschal_rc_3_20130717.rdf (consulted 28/10/13)
93. http://heritagedata.org/test/seneschal_rc_1_20130717.rdf (consulted 28/10/13)
94. http://heritagedata.org/test/seneschal_rc_10_20130717.rdf (consulted 28/10/13)
95. http://heritagedata.org/test/seneschal_rc_11_20130717.rdf (consulted 28/10/13)
96. http://purl.org/pico/thesaurus_4.1.0.skos.xml and
http://www.culturaitalia.it/opencms/export/sites/culturaitalia/attachments/thesaurus/4.3/thesaurus_4.3.0.skos.xml
97. <http://www.iccd.beniculturali.it/getFile.php?id=1691> (consulted 28/10/13)
98. <http://www.erfgoedthesaurus.nl/> and
http://www.cultureelerfgoed.nl/sites/default/files/u4/ABR_website2.pdf (consulted 1/11/13)
99. <http://kansalliskirjasto.onki.fi/rest/v1/ysa/data> and
<http://kansalliskirjasto.onki.fi/rest/v1/sapo/data> (consulted 1/11/13)
100. <http://museum.zib.de/museumsvokabular/> (consulted 23/11/13)
101. <http://archwort.dainst.org/thesaurus/de/vocab/> (consulted 23/11/13)
102. <http://www.getty.edu/research/tools/vocabularies/tgn/index.html> (consulted 21/10/13)
103. <http://www.geonames.org/> (consulted 5/10/2013)
104. <http://pleiades.stoa.org/downloads> (consulted 5/10/2013)
105. <http://pleiades.stoa.org/vocabularies/time-periods> (consulted 5/10/2013)
106. <http://logainm.ie/>
107. Bruce E. Bargmeyer and Daniel W. Gillman, Metadata Standards and Metadata Registries: An Overview,
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.119.5690&rep=rep1&type=pdf>
108. ISO 11179-1, Part 1 . ISO 11179-1, Part 1 - Framework for the Specification and Standardization of Data Elements
109. ISO 11179-3, Part 3 - Basic Attributes of Data Elements
110. <http://www.w3.org/TR/vocab-dcat/>