The Ariadne Registry of LORs

Jose Luis Santos, Joris Klerkx, and Erik Duval

joseluis.santos, joris.klerkx, erik.duval
@cs.kuleuven.be
Dept. Computerwetenschappen
K.U.Leuven
Celestijnenlaan 200A
B-3000 Leuven
Belgium

Abstract. The ARIADNE registry is one of the core components in an architecture that promotes interoperability of networks of repositories that facilitates the access to the learning content and encouraging the share and reuse of digital content. This paper explains the development of the ARIADNE registry of learning object repositories (LORs) and the role played by standards and specifications.

1 Introduction

One of the problems of Technology Enhanced Learning (TEL) is that creating Learning Objects (LOs) is expensive and time-consuming [1]. Sharing is one of the possibilities to address this problem. When content creators search for materials to reuse, they typically do not care about where the resource is located, but want to find the best quality materials that satisfy their needs [2].

LOs are typically stored in Learning Objects Repositories (LORs). In ARIADNE, considerable effort has been spent on the development of standards and specifications for LORs [3], including IEEE LOM [4], SQI [5], SPI [6], OAI-PMH [7] and PLQL [8]. These allow effective share and reuse of LOs between different LORs and networks of LORs.

One of the problems for managing all these repositories within a network is the scalability. Currently, every repository is added manuallyin the harvester or in the federated search layer service. However, this time consuming process requires that one person configures the targets and its parameters such as the requirements (e.g. query languages for querying or metadata formats for harvesting).

In addition, some extra information about the content inside of the repositories can be useful for managing this architecture. For instance, if we know that one repository contains LOs focused on mathematics, and we are interested in them, we can select this target to be harvested. To enable the sharing and exchange of this information with other networks of repositories, the information needs to be structured and managed.

This paper focuses on how the ARIADNE implements the ARIADNE Registry that has been integrated in the GLOBE architecture [9] and how ARIADNE addresses the problems described above by:

1. using a reference implementation developed for the ASPECT project [10] based on the concept of content collection. This content collection contains relevant information about the content of the collections and technical information for accessing them;
2. creating a registry for managing this information;
3. using standards and specifications for increasing the interoperability within networks of the repositories;
4. creating a network of registries in order to exchange the information between them and to be able to access the LOs of other networks. In this way, all the modifications done in a network can be widespread through all the networks of the registries.

The paper is organized as follows: section 2 which shows a possible use case of the approach. Section 3 introduces an explanation of the Registry Data Model co-developed in ASPECT project. Section 4 explains the Registry Architecture. Some statistics and data are presented in section 5.

2 Use Case. ARIADNE Foundation integrates the ARIADNE registry in its infrastructure

This section focuses on a use case for the ARIADNE registry and discusses: (i) the integration of the ARIADNE Registry in existing architectures and (ii) increasing the collaboration between different institutions for exposing their LOs. Also, we introduce some technical details about the implementation. The final goal is to explain the use of the registry in a non formal way.

ARIADNE Foundation has several repositories where the LOs are described by Learning Object Metadata. They expose the metadata using SQI and OAI-PMH. However, they have noticed that the number of repositories is increasing and it's difficult to manage all the information from other content providers. These content providers describe their LOs using different specifications like LOM and Dublin Core(DC), on the other hand some content providers expose the LOs using OAI-PMH and/or SQI. Looking for a good solution, they decide to integrate the ARIADNE Registry to manage this information because:

1. The ARIADNE Registry allows the definition of LORs using IMS LODE. That is open in terms of using specifications.
2. The ARIADNE Registry exposes the information using SQI, RSS and OAI-PMH. SQI allows to query the registry. RSS alerts subscribers when a new target is added. OAI-PMH allows to harvest all information from the ARIADNE Registry. These three specifications allow to integrate their ARIADNE Federated Search Layer which queries different repositories to obtain different LOs from different repositories, the ARIADNE harvester which

harvests metadata from different repositories and to build a federation of registries which allows to collaborate with different institutions.

The benefits that they obtain from the integration of the ARIADNE Registry are:

1. They have centralized all the information from all the repositories where they harvest from. It saves time in terms of administration tasks.
2. They can implement other services on top of the registry to check the availability of the different services. This service allows that other services can check this information before trying to access the targets. This information optimizes the performance behavior of these services.
3. The integration of the registry allows that new targets added are widespread by different network of LORs which are federated with ARIADNE network.

3 ARIADNE Registry Data model

The ASPECT project [10] has co-developed an application profile of IMS LODE[11]. It is based on IMS Dublin Core Collections Application Profile specification and complemented with ISO 2146 and IEEE LOM [12]. This specification uses the concept of content collections. A content collection is defined as a group of digital content which is exposed to the world through some protocols based on standards or specifications.

ARIADNE has chosen this specification, because it is not restrictive in terms of use of specifications, and it increases the possibilities of interoperatibility between architectures. In addition, the model does not restrict how the content collections are created. This is an important issue, because the content providers can choose how they create them and can offer metadata information about the collections that they are interested in.

This model is represented by a schema that contains three main elements Content Collection, Metadata Collection and Protocol.

1. Content Collection contains information about access rights, authoring, title description, keywords, etc. This is information about the content itself.
2. Metadata Collection contains information about how the metadata is exposed. Here, the content provider has to define which specification are used to expose metadata. This part of the specification has an element for defining the specification called Protocol Implementation Description which is used for defining extra information like the query language supported by an SQI interface or sets supported by an OAI-PMH interface.
3. Protocol contains specific information about the specification used like the URL of the schema, namespace or the binding location.

Several examples can be found at the ARIADNE Registry site [13].

4 ARIADNE Registry Architecture

The implementation of the registry enables ARIADNE to build a federation of registries, to provide access to collection information using SQI and to publish new content collections using SPI.

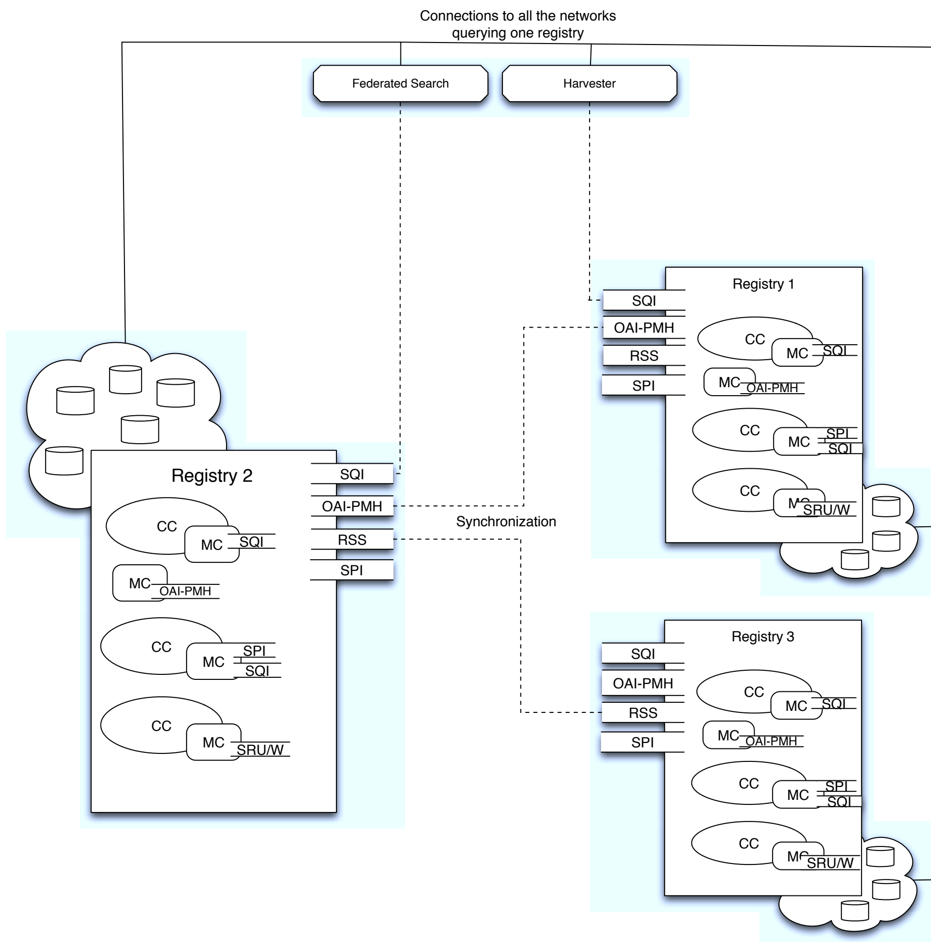

Fig. 1. ARIADNE architecture approach

A registry can contain metadata collections (MC in the figure 1) or content collections (CC). The latter contain metadata collections (MC).

The final goal of the registry is to create a network of networks of repositories, similar to current DNS functionality on the internet. All the registries are synchronized so if a harvester or a federated search service queries a registry, they can access all the content collections or metadata collections in the different networks. This approach is important: it allows automated discovery, decreases time spent managing repository information, and allows automated widespread updates.

This architecture shows how the registry can be synchronized with other registries. The registry exposes its contents using OAI-PMH so that other registries can harvest its content. In addition, the registry exposes its updates also using the RSS 2.0 specification, so that registries can be synchronized also with RSS feed readers. This RSS system can be used as a notification system, as it contains all the targets published.

The registry uses SQI for querying because it is neutral in terms of query language or results formats. Consequently, all SQI clients can query the registry. The registry supports different query languages like PLQL, VSQL and Lucene Query Language.

Finally, the registry implements an SPI interface for publishing content collections. For instance, when other content collections are harvested, they are inserted using SPI to publish content by reference, because the metadata in this case defines a content collection already published elsewhere.

One of the main advantages of this implementation is that ARIADNE Registry is based on the ARIADNE Repository. This software is a flexible implementation that allows different models of metadata. For instance, it supports LOM, ILOX, Dublin Core and this paper explains how it supports IMS LODE Registry specification. The idea behind this software is that we can index all kind of xml document. However, the final goal is that we can built combined queries, for instance, using PLQL, based on the content of different tags.

Trying to validate this architecture, we have integrated the ARIADNE harvester with the registry. The following sequencing diagram 2 shows how a system administrator can configure the harvester using the current implementation where the system administration can query the registry and add the targets that he/she is interested in.

5 Related Work

One of the main goals of ARIADNE is to encourage sharing and reusing LOs. For this purpose, we have implemented a registry using the concept of content collection, integrating some technical information in the data model and allowing the federation of registries. Table 1 shows the similarities between existing approaches and the ARIADNE approach.

Domain Name System (DNS)[14] has conceptual similarities with the approach explained in this paper. While DNS translates a name identifier to a number identifier, the registry translate abstract information contained by the content collection to a technical information contained in the metadata collection. In addition, the federation of DNS allows the propagation between them of the new sites. It is a similar approach that ARIADNE Registry implements with

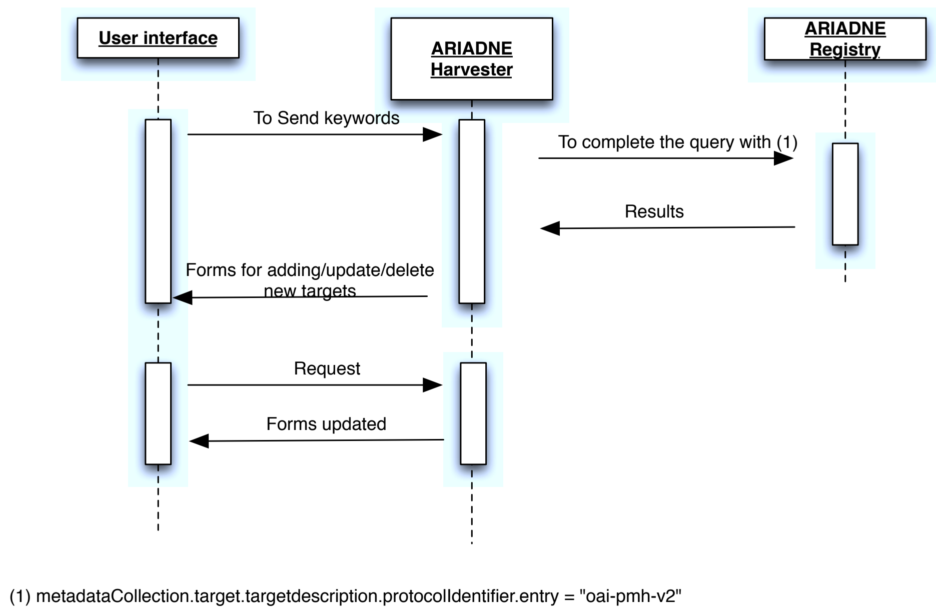

Fig. 2. Interaction diagram - Configuring the harvester

Table 1. Similarites with an existing approaches (Non-exisiting, Weak and Strong)

Different approaches	*Content collection*	*Extra technical information*	*Federation*
DNS	Non-exisiting	Non-exisiting	Strong
MWSDI	Weak	Strong	Weak
IESR/OCKHAM	Strong	Weak	Weak
CORDRA	Strong	Weak	Weak

the federation of the registries (the new content collections should be propagated among them)

METEOR's Web Service Discovery Infrastructure (MWSDI) uses UDDI specification for defining the interfaces (web services) and it defines an ontology for providing support to the infrastructure using OWL. MWSDI is based on the need to decentralize an architecture for several reasons such as geographical location, nature of registered services, business functionality, technical specifications and so on [15][16]. For this reason, it contains an ontology with details of the domain, Registries, Ontologies and Registries federation and network of relationship among them. The main difference is that MWSDI doesn't contain a content collection approach and the ARIADNE registry doesn't implement an hierarchical architecture of registries.

JISC Information Environment Service Registry (IESR)[17] and OCKHAM NSF / NSDL Registry [18] are developed under Global initiatives Registries [18]. They introduce the concept of content collection approach, while IESR uses Research Support Libraries Programme Collection Description schema (RLSPCD) and Dublin Core, which allows the definition of some technical details of the services which provide access to the collections[17], OCKHAM NSF/NSDL uses Dublin Core. They consider the definition of how the repositories expose the metadata considering protocols like Z39.50, Web Service SOAP, SRW and OAI-PMH. However, they don't provide extra information in their implementations such as sets supported by OAI-PMH or query languages. The registries expose their metadata through a SRU client [19, 16].

CORDRA / ADL registry (ADL-R) merges the concepts or MWSDI and IESR/OCKHAM, introduces a hierarchical structure of registries and the concept of content collection based on repositories which accept LOM as a metadata of LOs [20] .

At the end, we can conclude that Ariadne Registry incorporate the best characteristics from other approaches:

1. Conceptually, we implement similar approach to DNS widespread functionality.
2. IMS LODE Registry allows the definition of the services like UDDI specification concept do.
3. IMS LODE Registry allows the content collection approach like the above approaches mentioned do.
4. ARIADNE Registry allows the federation of the registries.

6 Statistics

In this section, we include some information about the metadata collections inserted and how they expose the metadata. It includes the number of records harvested using the OAI-PMH protocol. However, SQI or SRU/W don't provide a way to get all the possible results, thus the SQI and SRU/W targets do not contain information in date and records harvested field (table 2).

Table 2. Records harvested and protocols

id	*date*	*records*	*SQI*	*SRU*	*OAI*	*SPI*
dum_rvp_cz	09 Apr 2010	2323	No	No	Yes	No
atlas_fri_uni − lj_s	12 Apr 2010	4400	No	No	Yes	No
openlearn_open_ac_uk	12 Apr 2010	535	No	No	Yes	No
sodis_de	12 Apr 2010	2277	No	No	Yes	No
sdt_sulinet_hu	12 Apr 2010	18864	No	No	Yes	No
ait_opetaja_ee	12 Apr 2010	4050	No	No	Yes	No
193_43_17_27	12 Apr 2010	19309	No	No	Yes	No
193_93_132_187	12 Apr 2010	3541	No	No	Yes	No
www_yteach_com	12 Apr 2010	10	No	No	Yes	No
aspect	12 Apr 2010	124933	Yes	No	Yes	Yes
aspect_cup_cam_ac_uk	12 Apr 2010	8839	No	No	Yes	No
melt_contento_se	12 Apr 2010	3014	No	No	Yes	No
spindeln_iml_umu_se	12 Apr 2010	6699	No	No	Yes	No
lom_emokykla_lt_u_1	12 Apr 2010	1101	No	No	Yes	No
195_82_131_106	12 Apr 2010	196	No	No	Yes	No
aplitic_xtec_cat	12 Apr 2010	9876	No	No	Yes	No
www_klascement_net	12 Apr 2010	17739	No	No	Yes	No
www_portaldasescolas_pt	12 Apr 2010	1554	No	No	Yes	No
fenix_isftic_mepsyd_es	12 Apr 2010	26666	No	No	Yes	No
Mace_0	13 Apr 2010	180461	No	No	Yes	No
Globe	13 Apr 2010	385739	Yes	No	Yes	No
caad_asro_kuleuven_be	3 May 2010	450	No	No	Yes	No
caad_asro_kuleuven_be_0	3 May 2010	65	No	No	Yes	No
caad_asro_kuleuven_be_1	3 May 2010	171	No	No	Yes	No
caad_asro_kuleuven_be_2	3 May 2010	890	No	No	Yes	No
caad_asro_kuleuven_be_3	3 May 2010	2773	No	No	Yes	No
caad_asro_kuleuven_be_4	3 May 2010	50	No	No	Yes	No
caad_asro_kuleuven_be_5	3 May 2010	2106	No	No	Yes	No
cumincad_scix_net	3 May 2010	9636	No	No	Yes	No
ariadne_members	3 May 2010	3290	Yes	No	Yes	No
ariadne_partners	3 May 2010	71600	Yes	No	Yes	No
icoper	3 May 2010	16454	Yes	No	Yes	No
eun	3 May 2010	187722	Yes	No	Yes	No
ourj_code	3 May 2010	297	Yes	No	Yes	No
iskme	3 May 2010	30907	No	No	Yes	No
oai_rails_it_com	3 May 2010	211	No	No	Yes	No
oai_rails_it_com_0	3 May 2010	2223	No	No	Yes	No
oai_rails_it_com_1	3 May 2010	29	No	No	Yes	No
kocw	Not applicable	Not applicable	Yes	No	No	No
merlot	Not applicable	Not applicable	Yes	No	No	No
educanext	Not applicable	Not applicable	Yes	No	No	No
lornet	Not applicable	Not applicable	Yes	No	No	No
espol	Not applicable	Not applicable	Yes	No	No	No
agrega	Not applicable	Not applicable	Yes	No	No	No
nime	Not applicable	Not applicable	Yes	No	No	No
fedora	Not applicable	Not applicable	No	Yes	No	No
Total		1151000	14	1	38	1

Table 2 shows that the registry contains fourteen repositories which support SQI, one repository which supports SRU/W, twenty-five which support OAI-PMH and one which supports SPI. As the ARIADNE harvester relies on the registry, we know how many objects are exposed through OAI-PMH. The information is shown in the table.

The result is that we expose 1151000 metadata instances through the ARIADNE registry so far, although, we have to remark that some of them are duplicated, because some targets harvest other targets. But it means that we can expose a large amount of metadata instances specifying the protocol to access to them. All the harvesters which implement OAI-PMH can connect easily to the registry, using an SQI client for finding the targets and harvesting all the metadata from all the repositories described in the registry.

7 Conclusions and future work

We have developed a LOR registry, which enables us to set up a federation of registries. Different federations of repositories can easily obtain information from other federations through the registries. In this context, we will deploy the ARIADNE Registry in the ASPECT project, and we will evaluate the pros and cons of the federation.

In the scope of this work, we have considered some of the main specifications about querying, harvesting and publishing like SQI, OAI-PMH and SPI. We have spent some efforts to define extra information for the SRU/W. However, we have to validate this specification. Also, we would like to consider more specifications such as OKI OSID.

Now, we have integrated the ARIADNE harvester with the registry, in addition, we have to work on the Federate Search Layer to allow querying based on the content of the registries.

Acknowledgment The ARIADNE Registry has been funded partially by the ASPECT project: Adopting Standards and Specifications for Educational Content (Grant agreement number ECP-2007- EDU-417008) and by the CEN/ISSS workshop on Learning Technologies (WS-LT). Thanks also to our colleagues Ivana Bosnic and Katrien Verbert for helping us in this work.

References

1. Ochoa, X., Duval, E.: Measuring learning object reuse. In: EC-TEL '08: Proceedings of the 3rd European conference on Technology Enhanced Learning, Berlin, Heidelberg, Springer-Verlag (2008) 322–325
2. Ternier, S., Bosman, B., Duval, E., Metzger, L., Halm, M., Thorne, S., Kahn, J.: Connecting oki and sqi: One small piece of code, a giant leap for reusing learning objects. Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications (2006) 825–831

3. Duval, E., Forte, E., Cardinaels, K., Verhoeven, B., Van Durm, R., Hendrikx, K., Forte, M.W., Ebel, N., Macowicz, M., Warkentyne, K., Haenni, F.: The ariadne knowledge pool system. Commun. ACM **44**(5) (2001) 72–78
4. Neven, F., Duval, E.: Reusable learning objects: a survey of lom-based repositories. In: MULTIMEDIA '02: Proceedings of the tenth ACM international conference on Multimedia, New York, NY, USA, ACM (2002) 291–294
5. Bernd Simon, Daniel Olmedilla, N.S.e., ed.: A Simple Query Interface for Interoperable Learning Repositories. Volume 11-18., Chiba, Japan (2005)
6. Ternier, S., Massart, D., Van Assche, F., Smith, N., Simon, B., Duval, E.: A simple publishing interface for learning object repositories. Proc. World Conf. Educational Multimedia, Hypermedia, and Telecommunications, Assoc. for the Advancement of Computing in Education (2008) 1840–1845
7. Lagoze, C., Lagoze, C., Van de Sompel, H.: The open archives initiative: Building a low-barrier interoperability framework. IN PROCEEDINGS OF THE ACM/IEEE JOINT CONFERENCE ON DIGITAL LIBRARIES (2001) 54–62
8. Ternier, S., Massart, D., Campi, A., Guinea, S., Ceri, S., Duval, E.: Interoperability for searching learning object repositories: The prolearn query language. D-Lib Magazine **14**(1-2) (2008)
9. Klerkx, J., Vandeputte, B., Parra, G., Van Assche, F., Duval, E.: How to share and reuse learning resources: the ariadne experience. In: Proceedings ECTEL Conference 2010. Accepted for publication. (2010)
10. project, A.: http://www.aspect-project.org/ (2010)
11. LODE, I.: http://www.imsglobal.org/lode/spec/imslodev1p0bd.html (2010)
12. Massart, D., Smith, N., Tice, R.: Design of data model and architecture for a registry of learning object repositories and application profiles. deliverable 2.2 from aspect project. ASPECT PROJECT (2009)
13. Example Registry Instance, I.: http://ariadne.cs.kuleuven.be/ariadne-registry/search/showmetadata.jsp?query=metadatacollection.identifier.entry
14. Mockapetris, P., Dunlap, K.J.: Development of the domain name system. SIGCOMM Comput. Commun. Rev. **18**(4) (1988) 123–133
15. Sivashanmugam, K., Verma, K., Sheth, A.: Discovery of web services in a federated registry environment. In: ICWS '04: Proceedings of the IEEE International Conference on Web Services, Washington, DC, USA, IEEE Computer Society (2004) 270
16. Xiang, X., Morgan, E.: Exploiting 'light-weight' protocols and open source tools to implement digital library collections and services. D-Lib Magazine **12**(10) (2005)
17. Apps, A.: A registry of collections and their services: from metadata to implementation. In: DCMI '04: Proceedings of the 2004 international conference on Dublin Core and metadata applications, Dublin Core Metadata Initiative (2004) 1–10
18. Registries, G.I.: http://www.globalregistries.org/ (2010)
19. Morgan, E., Frumkin, J., Fox, E.: The ockham initiative - building component-based digital library services and collections. D-Lib Magazine **10**(11) (2004) 51–59
20. Jerez, H., Manepalli, M., Blanchi, C., Lannom, L.W.: Adl-r: The first instance of a cordra registry. D-Lib Magazine **12**(2) (2006)