
ECP-2005-CULT-038137



**TOWARDS AN ACCESSIBLE SCIENCE: FACILITATING
ACCESS TO SCIENTIFIC
DIGITAL RESOURCES FOR VISUALLY IMPAIRED
STUDENTS**

**D. Tools to Enable Communication
among University Library and
Distributed Digital Library
Implementation**

Deliverable number	<i>D4.3</i>
Dissemination level	<i>Public</i>
Delivery date	<i>26th February 2009</i>
Status	<i>Final</i>
Author(s)	<i>@Science Thematic Network Consortium</i>



eContentplus

This project is funded under the *eContentplus* programme¹, a multiannual Community programme to make digital content in Europe more accessible, usable and exploitable.

© Copyright 2006 – 2008, The *@Science* Consortium, consisting of:

- Università degli Studi di Milano
- Universität Linz, Institut “integriert studieren”
- Katholieke Universiteit Leuven
- Faculty of Mathematics Physics and Informatics, Comenius University
- Unione Italiana Ciechi – sezione provinciale di Verona
- Université Pierre et Marie Curie, Paris

This document may not be copied, reproduced, or modified in whole or in part for any purpose without written permission from the *@Science* Consortium. In addition to such written permission to copy, reproduce, or modify this document in whole or part, an acknowledgement of the authors of the document and all applicable portions of the copyright notice must be clearly referenced.

All rights reserved.

This document may change without notice.

¹ OJ L 79, 24.3.2005, p. 1.



EXECUTIVE SUMMARY

This document at first introduces some successful collaboration initiatives among libraries aimed to address the needs of blind and visually impaired people. Subsequently, it illustrates tools which can facilitate exchange of digital objects among digital libraries based on digital repositories. These tools are the Open Archives Initiative protocols and practices. Attention is drawn on accessibility statement and discovery of digital resources. In the last section the digital repository of the @science network is introduced.

TABLE OF CONTENTS

EXECUTIVE SUMMARY	3
1 INTRODUCTION	6
2 INTERLIBRARY INITIATIVES ADDRESSED TO SIGHT IMPAIRED	8
2.1 Introduction	8
2.2 Czech Republic	9
2.3 France	10
2.4 Ireland	10
2.5 Italy	10
2.6 United Kingdom	11
revealWeb	11
Open Rose Group	12
2.7 United States	13
2.8 Conclusions	13
3 OPEN ARCHIVES PRACTICES	15
3.1 Introduction	15
3.2 OAI-PMH	15
3.3 Metadata	16
3.3.1 Introduction	16
MODS: Metadata Object Description Standard	18
Qualified Dublin Core	18
MARCXML	19
ETD-MS: Electronic Theses, and Dissertations Metadata Standard	19
IMS Learning Resource Meta-data Specification	19
3.3.2 Metadata for Accessibility	20
3.4 OAI Repositories of Learning Objects: success experiences	21
ArXiv	22
Ariadne	22
EducaNext	22
JEM network	23
PlanetDR-URV	23
SMETE	23
3.5 Metadata Harvesters	23
4 THE @SCIENCE REPOSITORY	25
4.1 Introduction	25
4.2 Repository features	25
4.3 Implementation	26
5 REFERENCES	27





1 INTRODUCTION

In the last fifteen years, awareness of including blind and partially sighted people in academic and public library services has widely grown all over Europe. Libraries aim to make available to the public resources for reading. In particular, academic libraries aim to make available to students, researchers and academic staff a suitable selection of scholarly resources for teaching, learning and research purposes (e.g. books, journals, graduation and PhD theses, technical reports, etc.). Library services can become accessible to blind and partially sighted if at least some basic activities can be achieved by library users (for detailed information see [Access to libraries for persons with disabilities - CHECKLIST \[1\]](#)):

- reading resources can be searched, discovered and requested for reading through a procedure accessible to blind and partially sighted (e.g. an accessible catalogue on the web);
- reading resources can be delivered in accessible formats (e.g. as audio recordings, accessible electronic files, large print, etc.).

The first problem was tackled by making available accessible OPAC and MetaOPAC on the web. Basically that means to comply with web accessibility guidelines in developing the web user interface (further in [2] and [3]). The second problem is far more challenging. To face this problem, different strategies based on interlibrary collaboration have been put in place over the years:

- interlibrary loan of accessible material. Academic libraries make agreements with specialist organizations in order to provide students and academic staff with reading resources in accessible formats. Library patrons search the reading resource in the library catalogue and then the library is in charge of managing the request with the specialist organization which owns the resource in alternative format. This strategy is successful only if the academic library has access to a large catalogue of resources in accessible format owned by one or more specialist organizations;
- the academic library, often in collaboration with university support centres for disabled students, is able to adapt reading resources in accessible formats. This usually happens especially for learning material developed by university professors. These resources are often catalogued in separate catalogues;
- groups of academic libraries and specialist organizations are organized in consortia so that catalogues of resources in accessible formats are shared and interlibrary loan can be achieved.

All of these strategies are based on collaboration between different academic libraries and specialist organizations. That means in particular:

- making catalogues of resources in alternative formats accessible to all collaborating libraries or merging catalogues and set a procedure to keep the common catalogue updated. Unfortunately there is no widely accepted standard to catalogue resources in accessible formats therefore, different organizations have adopted over the years different conventions to catalogue resources in accessible formats;



-
- loan mechanisms of resources in accessible formats have to be set in compliance with copyright legislation. No relevant restrictions exist for loan of Braille printed sources or audiobooks in analog formats. Far more restrictions concern loan of material in digital format. It is therefore necessary to ask the permission to the right holder.

In section 2 the most successful experiences throughout Europe are illustrated.

A complementary approach in making available accessible content is through digital repositories. Over the years academic libraries have more and more undertaken projects to make available their digital contents through digital repositories. A repository is an online, searchable, web-accessible collection of digital objects (in case grouped in collections) which are developed in the academic context (e.g. works of research, learning objects such as lecture notes and slide presentations prepared by university professors for their courses, books made available in digital format, research papers, conference proceedings, journals, etc.). These repositories aim both to increase access to scientific content (e.g. to facilitate reading of research papers) and also to make available to the public (e.g. to students, researchers, etc.) reusable study material. Digital repositories are often built to serve community of users in a certain institution (they are known as institutional repositories). In other circumstances, repositories are built to serve scholars in a specific discipline (e.g. the arXiv repository for physics, mathematics and theoretical computer science). In the years many institutional repositories and discipline oriented repositories have been built by using several different conventions for metadata schemata, protocols for exchange of digital objects, admissible formats for files to be archived in the repository and submission and distribution policies. Thanks to the evolution of the open archive model, in particular since 2002 (Budapest Open Access Initiative [4]), standards for metadata and protocols for metadata harvesting have been developed and they have been getting more and more widespread all over the world in the design and implementation of digital repositories. Compliance with these standards will ensure better communication between those university libraries (or generally speaking content providers) which will maintain a digital repository of scholarly material. Section 3 will illustrate the basic open archives practices which can ensure better access, use, reuse and exploitation of scholarly content stored in digital repositories. Blind and partially sighted people who need access scientific content will widely get advantage of open access initiative practices. Some basic reasons account for that:

- open access initiative is making available more and more scientific content in open formats, which are far more adaptable for assistive technology;
- federated searches in multiple repositories gives the opportunity to discover related resources, which may be available in different formats. Therefore, for a visually impaired person, it is more likely to find a resource in accessible format concerning a certain topic;
- metadata harvesting by using the OAI-PMH protocol is independent of the user interface. Therefore, if the user interface to search, retrieve and display resources in repositories is not accessible for visually impaired, an accessible user interface which exploits OAI-PMH to harvest metadata can be developed by respecting W3C Web Content Accessibility Guidelines.

2 INTERLIBRARY INITIATIVES ADDRESSED TO SIGHT IMPAIRED

2.1 Introduction

In order to increase the availability of reading resources in accessible formats, many collaboration activities between public or academic libraries, libraries for the blind and specialist organizations have been carried out both all over Europe and outside Europe. Collaboration activities have been focused on:

- making available online catalogues of accessible resources available in multiple institutions (e.g. specialist organizations, libraries for the blind and also academic libraries);
- making available online procedures to request reading resources for lending.

In achieving these goals, many problems have been tackled by involved institutions:

- cataloguing resources in accessible format. There is no widely accepted standard to catalogue resources in accessible formats. Therefore, each institution has adopted over the years, local conventions. Therefore, unifying catalogues in one catalogue or accessing multiple catalogues are challenging activities, which often discourage collaboration;
- keeping updated catalogues. Since each institution has adopted different procedures to catalogue reading resources, new procedures to keep updated a shared catalogue have been often necessary, thus causing changes in the usual working activities. That has often led to the need of training;
- interlibrary lending procedures. New procedures have been designed to make possible interlibrary lending of resources in accessible format (e.g. audiobooks, Braille books, etc.). Basically two strategies have been adopted: delivery through surface mail or exchange of source files to be printed or recorded by the library requesting the source. The first strategy has been facilitated thanks to mail facilitations in sending reading resources for blind or partially sighted (e.g. in United Kingdom, Germany and Italy these mail services are at no cost). Exchange of files in digital formats is far more manageable, but copyright issues have to be faced in advance;
- copyright issues in lending reading resources in digital formats. In order for libraries to lend accessible reading resources in digital formats agreements with publishers and permission requests have to be put in place. Agreements have been sometimes reached with publishers at local level (e.g. for students with disability enrolled at a certain university) or at national level. Nonetheless, it is far more difficult to reach these agreements at transnational level. Actually, on the one hand publishers are not the right holders for distribution of published works outside a certain country, on the other hand publishers are worried about losing the control on the digital copy of the work.

Up to now these initiatives have been successful only in some countries. These success experiences can be considered as pilot stories for further evolution of these services. The following sections describe some of



the best collaboration interlibrary initiatives and collaboration activities between libraries and and specialist organizations in order to facilitate distribution of reading resources to sight impaired readers.

2.2 Czech Republic

The most advanced initiative in Czech Republic (and one the most advanced in Europe) to interconnect academic libraries and specialist organizations to the purpose of making available a union catalogue and interlibrary loan of accessible reading resources is The portal of university libraries for the blind [13] developed and maintained by Masaryk university in Brno [14], Tiresias Centre [15] in collaboration with Charles University Computer Centre, Tereza Centre at CVUT in Prague. The aim of the portal has been enabling sight impaired students to access a number of academic libraries and specialist centres which make available resources in accessible format by using one uniform web interface instead of many different tools. This portal makes possible browsing of library catalogues through Z39.50 protocol. The web interface is compliant with accessibility guidelines developed by W3C-WAI working group and it makes available a user interface especially designed for sight impaired, based on a 3 levels hierarchical structure for searching the catalogues and browsing the results. The search algorithm does not browse one catalogue locally stored, but it searches more catalogues belonging to each institution. The bibliographic record makes available many details about the resource in accessible format, anyway some slight differences may exist according to the library. In particular the following fields are available:

- title;
- author;
- order and place of publication, publisher, year of publishing;
- bibliographic data of the original publication (e.g. in standard print or as inaccessible electronic file);
- extent, series, notes, ISBN;
- information on the location of the electronic document. This field is available for the documents which are delivered in electronic format. If they are also available for download in electronic format, there is also a link to the download page in the digital library where the document is available. An authentication procedure is required before downloading the document.

At present the portal interconnects the following catalogues:

- Masaryk University Library for the Blind;
- Library of the Tereza Center at CVUT, Prague;
- Charles University in Prague Collected Catalogue;
- Masaryk University Collected Catalogue;



-
- Catalogue of the University Library at University in Hradec Kralove (a part of it is the digital library of Tyflopedic cabinet).

2.3 France

In France there are not initiatives for networking academic libraries to address the needs of sight impaired students. Instead, the association Braillet [16], founded in 1997, interconnects a number of libraries, specialist organizations and publishers in order to make available reading resources in accessible formats. Braillet maintains server Helene, which delivers digital sources of reading material to centres which produce them in Braille, large print, audio, etc. Server Helene was developed in partnership with INRIA Rhône-Alpes (INRIA: Institut national de recherche en informatique et en automatique) and it is supported by the Directorate of books and reading (Ministry of Culture and Communication). In 2008 the server included about 4000 titles, most of which were copyrighted works. Digital files are used by 67 centers specialized in producing material for sight impaired in France. Since 2006, Braillet association has managed Helene library. It is based on server Helene and it delivers reading resources in digital accessible formats to sight impaired people. Registration is free on presentation of a medical certificate. A digital rights management system is used to keep track of lent files. The catalogue includes resources from institutions collaborating with Braillet association. It is the only experience of lending of digital documents in France.

2.4 Ireland

In Ireland there are many associations which produce reading resources in alternative formats for the blind (e.g. National Council for the Blind of Ireland, National Braille Production Centre, etc.). In particular, many Irish universities contribute to adapt study material in accessible formats. All these contributions are accessible through the Textaccess Union Catalogue for the Print Disabled at Third Level. This union catalogue can be accessed via a web interface. At present there are 28 participating universities and institutes for technology. The resources are catalogued according to MARC21, with some extensions for resources in accessible formats. The TextAccess website also provides resources for librarians about accessibility issues.

2.5 Italy

In Italy there are some specialist organizations which produce most part of the reading resources in accessible formats. In particular: Biblioteca Italiana per i Ciechi Regina Margherita [17], Centro del Libro Parlato - Unione Italiana Ciechi [18], Associazione Libro Parlato Lions [19], Centro Internazionale del Libro Parlato [20] and many more initiatives, including university support centres and libraries which in some circumstances adapt learning material for sight impaired students. All of these initiatives have catalogued the resources produced in their own catalogues, by adopting many different catalogue conventions. Especially



organizations based on volunteers often developed in-house catalogue rules which do not refer to any cataloguing system. Since 2002 has been developed in Italy the Lettura Agevolata project [21]. The aim was to maintain a catalogue of resources in accessible formats available in many different catalogues of libraries and specialist organizations for sight impaired all over Italy. The project has been promoted by Comune di Venezia (Venice municipality) since its early beginning. A union catalogue, specifically designed to store information about resources in accessible formats (e.g. Braille books, documents in many digital formats, audiobooks, large print books, etc.) is archived and maintained updated by Lettura Agevolata staff. Each organizations periodically provides its own updates to Lettura AGEvolata staff, which are converted into the catalogue format of the union catalogue and loaded. Users can access the catalogue through a web interface tailored for sight impaired, which comply with web accessibility guidelines by W3C-WAI. At present, about 96000 resources in accessible format are available in the catalogue. Loan of reading resources in accessible formats depends on each association, library or specialist organization. Unfortunately not all of the Italian initiatives in making available accessible resources are catalogued by Lettura Agevolata project, especially initiatives led on local basis in universities and secondary schools. Another source of information about accessibility issues for libraries can be found in the CABI newsletter, which is delivered by Biblioteca Nazionale Marciana - Venezia(Marciana National Library in venice) [22].

2.6 United Kingdom

In the United Kingdom there are the best experiences in Europe of collaboration between libraries and specialist organizations in making available accessible reading resources to sight impaired readers.

revealWeb

Revealweb [5] for the first time in the United Kingdom has provided access to a union catalogue of reading resources in accessible formats for print impaired readers. Through Revealweb is possible to obtain information about the availability of a reading resources in one or more accessible format, to locate it (e.g. in a specialist library for the blind, in the catalogue of a specialist organization, etc.) and to request it for lending or to buy it. RevealWeb is the result of a number of joint efforts. In 1999 the Department of Culture, Media and Sport (DCMS) made a grant, to be administered by the Library and Information Commission and Share the Vision, to improve library and information services for visually impaired people.

As a first step, UKOLN [6], a centre of excellence in digital information management, providing advice and services to the library, information and cultural heritage communities, based at University of Bath,

was commissioned to design the technical specification of a union catalogue of resources in accessible formats. The already existing National Union Catalogue of Alternative Formats (NUCAF) was not directly accessible by end-users, the data it held was incomplete and sometimes incorrect; and its creator,

the Royal National Institute of the Blind [7] had not enough resources to maintain it. RNIB [7] and NLB [8] signed a partnership agreement to jointly develop a new online union catalogue of accessible resources.



UKOLN and GEAC Library Solutions [9], which had already undertaken a similar activity for the Canadian Institute of the Blind, were in charge of designing the catalogue and implementing the system. The first basic design choice concerned the catalogue. The bibliographic format and the content of records were established. MARC 21 Bibliographic Format [10] was chosen with some extensions. A new bibliographic record is inserted for each standard print item that has been used to create accessible versions. The details about all the accessible versions produced from the standard print item are attached to the main record. This approach facilitates the discovery of all the accessible formats available for a certain title. Some extensions were necessary to add information about accessible formats. Visually impaired people would need many details about a format in order to assess whether they can use an item. For example, Braille exists in three grades; in grade 1 each letter, number and punctuation mark has a Braille code. Grades 2 and 3 use contractions. Not all Braille readers are able to read grade 1,2 and 3 Braille. One more example concerns the format itself. For example, DAISY books can be available only as audio or as synchronized text and audio. According to the reader's preferences one or the other version may be more suitable (e.g. readers who became blind as adults usually prefer only the audio version). Another example concerns the need for a special reader. For example, file recorded in DAISY format can be used only through a DAISY player. The 007 field coding was not specific enough and additional codes were added. A few local fields were also added. In addition to material that has been produced, records are also created for material under development. This means that users can see, for example, that although an accessible version is not currently available, one is in production and expected to be available within a certain period. Access to the catalogue was then provided through an accessible web interface. Revealweb is also compliant with Z39.50. Therefore, another Z39.50 compliant service could search the Revealweb catalogue. Revealweb now has details of 112,528 titles, with 107 organisations on the register of suppliers, and the site receives around 30000 hits a month; with over 1500000 hits since its launch. Revealweb provides a highly value-added service for sight impaired persons both in the United Kingdom and elsewhere. Actually, through Revealweb a sight impaired person can find whether an accessible version of a certain item is available and request it either for lending or buying it. From outside the United Kingdom libraries which address the needs of sight impaired readers can check for the availability of items in accessible formats. They cannot receive them straightforwardly because of copyright issues. At first permission has to be negotiated with the publishers. This procedure can anyway be faster than adapting the reading resource from scratch.

Open Rose Group

The Open Rose Group is an effective example of collaboration among university libraries in addressing training of library staff about disability. It comprises library staff involved in supporting users with disabilities at eight Yorkshire universities in the United Kingdom. The participating institutions are: Bradford, Huddersfield, Hull, Leeds, Leeds Metropolitan, Sheffield, Sheffield Hallam and York. Since 2003, the group has produced audio-visual supporting material for use in libraries to assist staff training on disability issues. The students participating in the initiative are volunteers from Open Rose Group member institutions who each have one or more experiences in disabilities. They volunteered to talk about their experiences of using



their Library in relation to their disability. The training material for library staff can be requested at: openrose@email.com.

JISC Initiatives

The JISC (Joint Information Systems Committee) promotes several initiatives about libraries and accessibility issues. In particular through TechDis [11] and Libraries of the Future [12].

The JISC TechDis mission is to support the education sector in achieving greater accessibility and inclusion by stimulating innovation and providing expert advice and guidance on disability and technology. The TechDis website makes available many documents about case studies and success experiences in making accessible content available through academic libraries.

The Libraries of the Future initiative discusses issues concerning the evolution of libraries. In particular, are addressed questions such as what is the impact of repositories and open access on the delivery of library resources, what is the role of libraries in making more widely accessible scholarly content, and more. These topics are often discussed also with respect to accessibility issues for people with disability. On the website gives access to an online community of experts in library science (e.g. through blogs and RSS feeds).

2.7 United States

The most advanced initiative in interconnecting libraries and specialist organizations in making available to sight impaired accessible digital resources for lending is the National Library Service for the Blind and Physically Handicapped [23]. It makes available a comprehensive catalogue of accessible resources developed in dozens of specialist centres all over the United States and Canada. The catalogue archives items by using MARC21 with some extensions for specifying accessibility characteristics. The catalogue is based on Z39.50 protocol. The participating centres make available their catalogues to be searched so that the results can be found in the union catalogue. The union catalogue is accessible through a web interface, which is tailored for different profiles. Reading resources are available for lending or purchasing only for sight impaired people from United States and Canada. This is the largest catalogue of accessible resources in the world. The NLS service at the Library of Congress makes available also a large selection of newsletters [24] which deal with accessibility issues in libraries.

2.8 Conclusions

The interlibrary services introduced in the previous sections are the most advanced all over Europe and outside Europe especially as for making available information about how and where resources in accessible formats can be obtained. As illustrated, many differences exist among these services, which prevent from achieving a transnational library service for sight impaired. First of all, resources in accessible formats can be successfully catalogued by extending some of the most widespread cataloguing formats (e.g. MARC 21 as in



UK, United States and Czech Republic). There is anyway no shared set of conventions to extend these formats. Moreover, some initiatives adopted a unique catalogue instead of searching distributed catalogues which are more efficiently updated by local organizations. Notably, these initiatives cover only a small part of the EU Member States and in some circumstances (e.g. in France and Italy), only some of the specialist organizations and libraries producing resources in accessible formats contribute to the union catalogue. One more relevant problem concerns delivery for interlibrary lending of resources in accessible formats. Lending has to be negotiated with publishers, which usually are due to make agreements for national distribution. In many circumstances, publishers themselves do not own the rights to distribute a certain publication in a different country. Only Brailenet association up to now could negotiate with some publishers licences for lending resources which mention a clause "for the whole world". A first step towards transnational exchange of resources among academic libraries can be taken as follows:

- discuss and adopt a well-defined library accessibility policy which includes cataloguing and lending of resources in accessible formats. The Office for Information Technology Policy (OITP) of the American Library Association (ALA) has published advice on accessibility policy for libraries, based on work by the University of Washington's DO-IT initiative [25];
- contact representatives of national initiatives for union catalogues of resources in accessible formats;
- adopt a cataloguing format for accessible resources which can be easily converted in the one used by the national initiative for union catalogue;
- generate a local catalogue of resources in accessible format produced by the local library;
- transfer on a periodical basis the catalogue to the national one or make the catalogue accessible (e.g. through Z39.50 protocol) by software agents;
- for each acquisition, negotiate with publishers the authorization for lending the resource in accessible format for sight impaired students to the maximum extent. To this purpose, reference to other existing experiences may be of help.



3 OPEN ARCHIVES PRACTICES

3.1 Introduction

The open access model has developed much through years. In 1990 the open access arXiv for physics papers was set up. In 1995 two remarkable open access initiatives were launched: WoPEc for economics, NCSTRL for computer science. In 1996 PubMed included Medline as open access archive. In 2001 the Public Library of Science (PLOS) was founded. In 2002 the Budapest Open Access Initiative was issued. In 2003 was released the Directory of Open Access Journals. In 2006 the European Commission issued a report in order to recommend open access for public research. In 2006, in the United States was approved the Federal Research Public Access Act of 2006 to recommend open access in federal research. In the same year at CERN was launched SCOAP3 to convert to open access particle physics

journals. In 2007, the European Research Advisory Board issued a recommendation in using the open access model for European Union funded research. Also the European University Association recommended open access for collecting research results. The European Research Council issued the ERC Scientific Council Guidelines for Open Access which requires the deposit of papers and scientific data in open access repositories. In 2007, 427 scientific societies are publishing 496

scientific journal based on open access model. All over the world, graduation works and PhD theses are more and more being archived in open access repositories.

Since 2002 also best practices and standards have been getting widespread, thus ensuring better access, use, reuse and exploitation of digital objects stored in repositories. These practices and protocols will be described in the following sections.

The open archives model is based on two classes of participants:

- **Data Providers.** Data providers archive resources in digital repositories, describe resources through metadata sets and make available metadata to harvesters through the OAI-PMH protocol (Open Archives Initiative Protocol for Metadata Harvesting [26]). Data providers can be, for example, university libraries which make available their content in digital format in a repository;
- **Service Providers.** Service providers harvest metadata from Data Providers through the OAI-PMH protocol in order to make available value-added services. Service providers can be for example search engines specialized for a certain knowledge domain.

3.2 OAI-PMH

OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting [27]), is the protocol which Data Providers and Service Providers must comply with in order to make available content and achieve metadata harvesting. This section introduces briefly the characteristics of the protocol:



-
- OAI-PMH requests are expressed as HTTP requests (through HTTP GET or HTTP POST methods). A typical implementation uses a standard web server that is configured to dispatch OAI-PMH requests to the software handling these requests;
 - all responses to OAI-PMH requests must be well-formed XML instance documents. Encoding of the XML must use the UTF-8 representation of Unicode. Character references, rather than entity references, must be used. Character references allow XML responses to be treated as stand-alone documents that can be manipulated without dependency on entity declarations external to the document;
 - dates and times are uniformly encoded using ISO8601 and are expressed in UTC throughout the protocol;
 - OAI-PMH supports the dissemination of records in multiple metadata formats from a repository. All metadata available from a repository can be required through the ListMetadataFormats request. For purposes of interoperability, repositories must disseminate Simple Dublin Core metadata. Further metadata formats can be added;
 - OAI-PMH makes possible to retrieve lists of records, sets or identifiers through list requests.

3.3 Metadata

3.3.1 Introduction

Metadata formats are groups of defined elements, or fields that can be used to describe a resource (e.g. administrative information, technical information, etc.). Making metadata sharable is the basic characteristic a data provider should implement. The most significant advantage to creating shareable metadata is that metadata will be interoperable, or meaningful when combined with metadata from other sources. By using metadata schemata and rules for creating metadata values similar to those used by others, the resources collected in a certain repository can meaningfully appear in search results together with related resources from other repositories. Therefore, the resources will receive more exposure, and end-users will have the opportunity to make previously unseen connections between related resources stored in different repositories. Data providers do not need know in advance which are the repositories which contain related resources. If metadata are sharable, the metadata harvesters will be able to find relations between resources.

Based on the book "Metadata in Practice" [28], seven characteristics must be satisfied to have good metadata:

- **completeness.** That means in particular choosing an element set which allows the resources to be described as completely as is economically feasible, and applying that element set as completely as possible;



-
- **accuracy.** The metadata must be correct and factual, and conforming to syntax of the element set in use;
 - **provenance.** Here provenance refers to the provision of information about the expertise of the persons creating the original metadata, and its transformation history;
 - **conformance to expectations.** Metadata elements, use of controlled vocabularies, and robustness should match the expectations of a particular community. Since the resource can be retrieved by metadata harvesters, the expectations are not often known in advance in any circumstance;
 - **logical consistency and coherence.** This characteristic is defined as element usage matching standard definitions, and consistent application of these elements;
 - **timeliness.** That refers in particular to keeping up-to-date metadata concerning resources which change over the time and making available metadata describing actually available resources;
 - **accessibility.** It refers to proper association of metadata with the resource it describes and readability by target users.

Metadata may be of high quality within a certain local context, not in a broader context. There are other characteristics which make quality metadata also sharable in a shared environment:

- **proper context.** In a shared environment, metadata records will become separated from any high-level context applying to all records in a group, and from other records presented together in a local environment. It is therefore essential that each record contains the context necessary for understanding the resource the record describes, without relying on outside information;
- **content coherence.** Enough information must be contained in metadata records in order to retrieve resources even if outside the environment the resource belongs to;
- **use of standard vocabularies.** The use of standard vocabularies enables the better integration of metadata records from one source with records from other sources;
- **consistency.** All decisions made about application of elements, syntax of metadata values, and usage of controlled vocabularies, should be consistent within an identifiable set of metadata records so those using this metadata can apply any necessary transformation steps without having to process inconsistencies within such a set;
- **technical conformance.** Metadata should conform to the specified XML schemata and should be properly encoded.

The Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH) enables service providers to harvest metadata exposed by data providers. Metadata harvesting through OAI-PMH makes possible all the



advantages of open access repositories. Therefore, for a repository to be OAI-PMH compliant is an indispensable characteristic. To this purpose:

- a Simple Dublin Core metadata record (Dublin Core Metadata Element Set 1.1 [29]) must be available for every item in the repository. Therefore, the Open Archives Initiative makes available an XML schema at: http://www.openarchives.org/OAI/2.0/oai_dc.xsd for simple Dublin Core, and reserves the metadata prefix oai_dc for this schema;
- in addition to the Simple Dublin Core record, it is advisable to provide alternative or richer metadata records in further metadata formats. Actually, Simple Dublin Core is not able to describe details of many resources, which could be present in a certain repository. Moreover, Simple Dublin Core does not convey information about the use of controlled vocabularies or encoding schemes used. repository. By supplying additional metadata formats which have the semantic richness to more clearly express meaning, data providers can help service providers make better use of their metadata. Metadata formats used must have an XML schema available for validation (see <http://www.openarchives.org/OAI/openarchivesprotocol.html#MetadataNamespaces>).

Some of the most used metadata formats are listed below. For a complete list, consult the Distinct Metadata Schemas report maintained by University of Illinois at Urbana-Champaign (on the web at: <http://gita.granger.uiuc.edu/registry/ListSchemas.asp>).

MODS: Metadata Object Description Standard

MODS may be a good choice for data providers which:

- locally depend of metadata definition practices bound to resource description standards commonly used in libraries;
- address to a community well-versed in library descriptive practices, yet also want robust records in a format accessible to service providers outside the core library community.

The MODS v.3.0 XML Schema is available at

<http://www.loc.gov/standards/mods/v3/mods-3-0.xsd>

Qualified Dublin Core

Qualified Dublin Core may be a good choice for data providers which:

- have a need for more granularity of description than is available in simple Dublin Core but not a fundamentally different approach to resource description as the library community may wish;



-
- use controlled vocabularies to be specified within the metadata records;
 - address many different communities which may have their own metadata practices.

A XML schema for Qualified Dublin Core can be created through the importation of the necessary namespaces and schemas:

<http://dublincore.org/schemas/xmls/qdc/2003/04/02/dc.xsd>

<http://dublincore.org/schemas/xmls/qdc/2003/04/02/dcterms.xsd>

<http://dublincore.org/schemas/xmls/qdc/2003/04/02/dcmitype.xsd>

MARCXML

MARCXML may be a good choice for data providers which:

- locally describe resources in MARC (e.g. many libraries);
- address mainly the library community.

The MARCXML XML Schema is available at

<http://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd>

ETD-MS: Electronic Theses, and Dissertations Metadata Standard

ETD-MS may be a good choice for data providers which:

- mainly archive theses (e.g. graduation or PhD theses) and dissertations in digital format;
- aim to contribute to aggregations of electronic theses and dissertations such as the Networked Digital Library of Theses and Dissertations (on the web at: <http://www.ndltd.org>).

The ETD-MS XML Schema is available at

<http://www.ndltd.org/standards/metadata/etdms/1.0/etdms.xsd>

IMS Learning Resource Meta-data Specification

It may be a good choice for data providers which:

- archive different sorts of learning objects;

-
- aim to make available learning objects to the educational community.

The XML schema can be found at:

http://www.imsglobal.org/xsd/imsmd_v1p2p4.xsd

protocols is the AccessForAll Metadata developed by IMS [30].

3.3.2 Metadata for Accessibility

In order to make available value-added services based on search and retrieval of accessible learning resources in digital format, a set of metadata which describes accessibility characteristics of resources is necessary. Current sets of metadata are not sufficiently specific to describe accessibility features of resources archived in a repository. The IMS Global Learning Consortium's Access For All Metadata Specification (AccMD) [30] enables recording and sharing information about the accessibility of resources archived in a repository. The metadata specification is designed to work in combination with a record of the user's needs and preferences as expressed in the IMS Accessibility for Learner Information Profile Specification (AccLIP) [31]. Together these are known as the Access For All specifications. The adoption of these specifications in describing resources and user preferences would highly facilitate retrieval of resources which fit certain accessibility requirements and exchange of these resources among different repositories. The Access For All metadata specification groups resources into two possible categories: primary resources and equivalent alternative resources. The primary resource is the initial or default resource which do not necessarily fits any particular accessibility characteristic. An equivalent resource is referenced by a primary resource. It is the adaptation of the primary resource in accessible format. All the semantic properties of the primary resource are meant to be preserved. Since the authors of most primary resources (and the authors of their metadata) are not likely to be informed about accessibility considerations and may not be motivated to create additional metadata, the metadata on the primary resources is kept to a minimum. Metadata regarding a primary resource simply describes what access modes are needed to use the resource (e.g., visual, auditory, text, etc.), whether the display and method of control is flexible, and referenced equivalent resources, if there exist some. Unlike authors of primary resources, it is anticipated that authors of equivalent resources are likely to be both motivated and informed about accessibility considerations. The description of the equivalent resource is therefore much more detailed and closely matches the AccLIP specification. This permits resource portals to search for equivalent content that overcomes a mismatch between the user's required sensory modalities and the modalities included in a primary resource. Integrating the accessibility specifications into digital libraries will require a number of coordinated efforts. Tools are being developed for authoring user profiles and resource metadata, or authors may be able to integrate the new data into existing personal preferences and resource metadata tools. The digital library service needs offer users the ability to record a set of accessibility preferences that can influence both the browsing interface and the way resources are located for the user. Finally, accessibility metadata must be added to the metadata record for each resource, to the extent possible. The



greater the amount of metadata available, the more powerful the system becomes. Tools for authoring and using AccLIP and AccMD are emerging. A well-developed example (even an open-source system) for an accessible digital library site is the Collection Workflow Integration System (CWIS) from the Internet Scout Project [32]. CWIS offers a preferences editor for AccLIP data and transforms the library interface to match each user's preferences. CWIS can also edit and store AccMD data about resources and use them to highlight resources according to a user's AccLIP preferences.

A different or complementary approach in describing accessibility characteristics of digital resources is the Accessibility Passport [33] being developed at JISC-TechDis. This idea is based on the assumption that accessibility is not only a matter of standards, but it's also about how materials get delivered and their fitness for purpose. The Passport is an online document (e.g. a form or a wiki on a website), that is available to everyone and the link to it is carried within the learning object. All those involved in the writing, sharing and delivering of learning materials have responsibilities for ensuring that all readers are not excluded from the end result. The way that materials are delivered can have a drastic effect on the range of users who can benefit, and many different individuals may have an impact on the final experience. Those who design and write software or learning materials have a high level of responsibility for the accessibility of their output, but currently there is no standard means for them to inform others about the way they intended the materials to be used. An interactive 'passport generator' prompts the commissioner/developer to define the accessibility features they have built in and provides links to good practice signposts. In the process, the input from the developer creates a user oriented summary of the accessibility opportunities and constraints in the materials in a document which travels with each learning object. The Passport carries information and suggestions from the author and designer about:

- Inbuilt accessibility;
- Alternate strategies and delivery methods to suit a range of contexts and learners.

It encourages feedback and incorporates suggestions from users due to being an editable wiki. Uniquely, the wiki format would stimulate all those involved in the building and delivering of materials to be involved in an ongoing analysis of the accessibility and effectiveness of the materials in practice, for the life of the object. This approach is being evaluated at JISC-TechDis.

3.4 OAI Repositories of Learning Objects: success experiences

This section introduces some repositories of digital learning objects compliant with the Open Archives Initiative specifications.



ArXiv

Website: <http://www.arxiv.org>

ArXiv open access repository makes available more than 520000 e-prints in Physics, Mathematics, Computer Science, Quantitative Biology, Quantitative Finance and Statistics. The contents of arXiv conform to Cornell University academic standards. arXiv is owned, operated and funded by Cornell University, a private not-for-profit educational institution. arXiv is also partially funded by the National Science Foundation. Metadata about papers is made available through the OAI-PMH protocol.

Ariadne

Website: <http://www.ariadne-eu.org>

Ariadne is an European association for sharing knowledge and fostering international cooperation in teaching that is open to the world. Ariadne was initiated in 1996 by the European Commission's telematics for education and training program. Since then, an infrastructure has been developed in Belgium and Switzerland for the production of reusable learning content, including distributed storage and discovery, as well as its exploitation in structured courses. The core of this infrastructure is a distributed library of digital, reusable educational components called the Knowledge Pool System (KPS) now actively used in both academic and corporate contexts. End users interact with the KPS through client tools. Java and web applications allow users to insert documents and their associated metadata into the KPS, search for relevant documents, and download them from the KPS.

EducaNext

Website: www.educanext.org

EducaNext is a service supporting the creation and sharing of knowledge for Higher Education. It is open to any member of the academic or research community. EducaNext fosters collaboration among educators and researchers, allowing to participate in Knowledge Communities, to communicate with experts in a certain knowledge domain, to exchange Learning Resources, to work together on the production of Educational Material: Textbooks, lecture notes, case studies, simulations, etc., to deliver distributed Educational Activities: lectures, courses, workshops, case study discussions, etc., to distribute electronic content under license.



JEM network

Website: www.jem-thematic.net

The JEM network is a European thematic network for Joining Educational Mathematics. It started on 1 August 2006. The goal of the JEM thematic network is to pool together the required expertise and to contribute to the coordination of content enrichment activities in the area of mathematics, to the maintenance of agreed standards and to the delivery of powerful synoptic high-quality user information and support pages, invoked in e-learning platforms operated by the partners. These activities will have tremendous long-term benefits for the quality of e-learning in mathematics. JEM network makes available a rich repository of digital learning objects about mathematics and science.

PlanetDR-URV

Website: <http://planet.urv.es/planetdr/>

PlanetDR is an open source Content repository developed under the umbrella of the spanish research project PLANET (Augmented Collaborative Plataform for content distribution and advanced training). PlanetDR supports standards such as IEEE LOM, IMS Content Packaging (IMS CP), IMS DRI (EduSource Communication Language-ECL) and LORI SQL. PlanetDR also provides a federation mode between PlanetDR servers and plans to support the FIRE protocol.

SMETE

Website: <http://www.smete.org/>

The SMETE Digital Library is a dynamic online library and portal of services by the SMETE Open Federation for teachers and students. SMETE makes available a wealth of teaching and learning materials about science, math, engineering and technology which are suitable for many ages.

3.5 Metadata Harvesters

OAI-PMH metadata harvesters are software applications which can retrieve digital resources by analyzing metadata attached through OAI-PMH protocol. At present a comprehensive metadata harvester is Oaister [34]. Oaister is a union catalog of digital resources. It provides access to these digital resources by harvesting their descriptive metadata records using OAI-PMH. Oaister currently provides access to 20220634 records from 1082 contributors. The web interface made available by Oaister is highly accessible for sight impaired



users. Retrieved resources are sometimes accessible and sometimes they are inaccessible. However, they are often available in more than one digital format (e.g. LaTeX files, Xhtml files, PDF and PostScript files) so that they can be adapted more easily in accessible formats. A reference to Oaister in the website of an academic library would be a valuable resource because it opens up access opportunities to many resources in many digital repositories. Furthermore, if digital repositories managed by academic libraries could be searched by Oaister, a larger public would be able to find the digital resources archived.



4 THE @SCIENCE REPOSITORY

4.1 Introduction

The @science thematica network aims at making available through a digital repository:

- learning objects about scientific university courses accessible to blind and partially sighted students;
- learning objects about guidelines, best practices, case studies concerning how to enable blind and partially sighted students to access scientific subjects.

Furthermore, the @science thematic network also aims at collecting contributions from institutions willing to share their learning objects. In order to achieve these goals a digital repository based on OAI practices and standards was integrated in the @science website. This repository can archive learning objects made available by @science participants, it can import learning objects collections from other repositories, it can expose metadata for harvesters through the OAI-PMH protocol. Since at the moment there is no automatic accessibility validator for many file formats which can be used to make available accessible scientific content, contributions, before being archived in the repository are checked by @science members to match accessibility characteristics.

4.2 Repository features

The basic features of the repository follow:

- learning objects are archived according to the metadata schema IMS LOM (IMS Learning Resource Meta-data Specification v1.3, available at: <http://www.imsglobal.org/metadata/>). This metadata specification turns out to be especially suitable for learning objects. Also, it is widely used by many repositories making available learning objects (see section 2.4). In particular it is used by the Joining Educational Mathematics network. This metadata specification makes available fields to specify intellectual property rights for learning objects. The @science repository will make available not only free learning objects (e.g. learning objects which are licensed under Creative Commons share-alike license), but also learning objects which cannot be distributed to everyone because of copyright restrictions. This feature will be managed by comparing user profiles, namely user permissions, with the IPR metadata field for each learning object. For example, according to national transpositions of EUCD directive about copyright (see Deliverable 4.1) some special conditions apply for people with disability. If a registered member is registered as visually impaired and makes available to one of the @science members a legal declaration of disability (e.g. a copy of certificate of disability), some learning objects could be accessible and could be downloaded by this registered user, not by all users;
- OAI-PMH Data Provider support. The repository can expose metadata about learning objects to be harvested by OAI-PMH2 compliant metadata harvesters (e.g. OAIster: www.oaister.org/);



-
- classification of the learning resources on basis of LivingTaxonomy for Mathematics. The Living Taxonomy Project is a collaborative effort aimed at creating a global set of open source, standards-based taxonomies for education. The purpose of these taxonomies will be to provide a free cataloging structure for the collection and sharing of education materials around the world. (further information at: <http://www.livingtaxonomy.org/>);
 - LOM records can be imported both from IMS LOM v1.2 and IEEE LOM v1.0 formats. That facilitates importing of already available learning objects which can be shared with other repositories;
 - LOM records can be exported in LOM XML to be exposed to harvesters or to be imported in other repositories;
 - metadata records can be edited either through an accessible XHTML form or through a separate module;
 - the system automatically detects the format of the record on export;
 - real time download counter on basis of exports, available at the node view page;
 - terms in the vocabularies are linked to views which show all the records that contain the same values in the field;
 - author names are linked to user profiles if the author is a registered member.

4.3 Implementation

The repository system integrated in the website @science is the one used by the Joining Educational Mathematics network []. This repository is open source, it can be integrated in websites developed by Drupal 5 and it is based on PHP 5.1 or later. This repository specifically addresses the needs of a community which works with learning objects of scientific content.

5 REFERENCES

1. Access to libraries for persons with disabilities – CHECKLIST
www.ifla.org/VII/s9/nd1/iflapr-89e.pdf
2. Web Accessibility: Practical Advice for the Library and Information Professional, edited by Jenny Craven, Facet Publishing, 2008
3. Web Accessibility Initiative
www.w3.org/wai
4. Budapest Open Access Initiative
<http://www.soros.org/openaccess/read.shtml>
5. Revealweb
<http://www.revealweb.org.uk>
6. UKOLN at University of Bath
<http://www.ukoln.ac.uk/>
7. Royal National Institute for the Blind
<http://www.rnib.ac.uk>
8. National Library for the Blind (NLB)
<http://www.gateway-uk.org/nlb.htm>
9. GEAC Library Solutions
<http://vubis-smart.com/html/homeeng.htm>

10. MARC formats

<http://www.loc.gov/marc/>

11. JISC - TechDis

<http://www.techdis.ac.uk/>

12. Libraries of the Future

<http://www.jisc.ac.uk/librariesofthefuture>

13. The Portal of University Libraries for the Blind

<http://www.teiresias.muni.cz/portal>

14. Masaryk University in Brno (CZ)

<http://www.muni.cz>

15. Teiresias Centre at Masaryk University in Brno (CZ)

<http://www.teiresias.muni.cz>

16. BrailleNet

<http://www.brailenet.org>

17. Biblioteca Italiana per i Ciechi

<http://www.bibciechi.it>

18. Centro del Libro Parlato - Unione Italiana Ciechi

<http://www.uiciechi.it/servizi/lp/introlp.asp>

19. Associazione Libro Parlato Lions

<http://www.libroparlatolions.it/>



20. Centro Internazionale del Libro Parlato

<http://www.libroparlato.org>

21. Lettura Agevolata

<http://www.letturagevolata.it>

22. CABI Newsletter

<http://marciana.venezia.sbn.it/catalogazione.php?ss4=23>

23. National Library Service for the Blind and Physically Handicapped (USA)

<http://www.loc.gov/nls>

24. NLS Newsletters

<http://www.loc.gov/nls/newsletters/index.html>

25. American Library Association (2007) Creating a Library Accessibility Policy,

<http://www.ala.org/ala/washoff/oitp/emaitutorials/accessibilitya/22.cfm>.

26. Open Archives Initiative Protocol for Metadata Harvesting

<http://www.openarchives.org/OAI/openarchivesprotocol.html>

27. Open Archives Protocol for Metadata Harvesting

<http://www.openarchives.org/OAI/openarchivesprotocol.html>

28. Metadata in Practice, ed. Diane I. Hillmann and Elaine L. Westbrook, Chicago: American Library Association, 2004, Thomas R. Bruce and Diane I. Hillmann

29. Dublin Core Metadata Element Set 1.1

<http://dublincore.org/documents/dces/>



30. IMS Access For All Metadata Specification

http://www.imsglobal.org/accessibility/accmdv1p0/imsaccmd_bestv1p0.html

31. IMS Accessibility for Learner Information Profile Specification (AcclIP)

http://www.imsglobal.org/accessibility/acclipv1p0/imsacclip_bestv1p0.html

32. Internet Scout Project

<http://scout.wisc.edu/>

33. Accessibility Standards Are Not Always Enough: The Development of the Accessibility Passport, Simon Ball, John Sewell, in proceedings of ICCHP 2008, LCNS, 5105, Springer

34. Oaister

<http://www.oaister.org>

35. A LOM repository with Drupal

<http://www.jem-thematic.net/en/node/716>).