European Commission
Directorate General XIII - E/4

Electronic publishing and libraries

Telematics for Libraries

METADATA WORKSHOP

REPORT OF THE WORKSHOP HELD IN

LUXEMBOURG, 1 AND 2 DECEMBER, 1997

FEBRUARY 1998

TABLE OF CONTENTS

Executive summary

On 1 and 2 December 1997, a workshop organised by the European Commission DGXIII/E4 took place in Luxembourg, on the subject of metadata.

The workshop consisted of a *Tutorial* led by the UK Office for Library and Information Networking (UKOLN), presentations of current projects in Scandinavia and the UK, and three breakout sessions on *Metadata creation*, *Harvesting*, and *Retrieval*.

Around 60 participants from organisations all around Europe attended the workshop, indicating a wide interest in the subject.

The workshop participants recognised that metadata standards are necessary, in addition to sector- or subject-specific description mechanisms, to ensure interoperability in resource discovery on the Internet.

The workshop concentrated on the emerging and currently the best-developed metadata format known as Dublin Core. It is however recognised that it concentrates fully on resource discovery and does not cover other requirements, e.g. for resource management or access restrictions.

The workshop, although recognising the usefulness of Dublin Core as a starting point in metadata descriptive standards, brought forward a number of concerns regarding the current state and the further development of Dublin Core:

- There is currently no formal responsibility for the maintenance of Dublin Core: development takes place in an informal group of invited experts which meets once or twice per year in what is known as the Dublin Core Workshop Series.
- The current technical state of Dublin Core is unstable: during the meetings of the Dublin Core group, changes are being made to the format and there is no convergence to a stable version.
- The use of the current Dublin Core metadata format is not supported by the existence of guidelines: some of the philosophy and terminology of Dublin Core is not obvious to the uninitiated user which could lead to different interpretations adversely affecting interoperability.

It was also identified that the current take-up of Dublin Core is slow and that there is a lack of critical mass. This seems to be a classical chicken-and-egg situation: authors and publishers do not invest in providing Dublin Core metadata if the Internet indexing services (the 'harvesters') do not utilise it, and harvesters do not collect Dublin Core and use it for selective indexing if there is not enough data available. If this situation cannot be changed, Dublin Core might not turn into reality.

The workshop identified a number of actions that could be taken to promote and encourage the use of Dublin Core, including the following:

1. There needs to be clarity about version control and maintenance of Dublin Core. The Dublin Core group, addressed through the mailing list META2, will be asked to give a clear statement about this.

2. Further pilot projects should be started to further develop experience, test out the issues and help realise a critical mass of Dublin Core metadata. The European Commission and national bodies like National Libraries might have a role to play by encouraging the provision of Dublin Core metadata in documents, e.g. in project deliverables and electronic documents in the national deposit.

3. The interest and requirements existing in Europe warrant the establishment of a European group of implementers discussing the practical issues of implementing metadata in general and Dublin Core in particular. The Luxembourg workshops, such as this December 1997 one and a second one scheduled for mid-1998, could develop into a regular series.

4. The liaison with other groups concerned with metadata, such as the CEN/ISSS working group on Metadata for Multimedia Information (MMI), should be established to ensure applicability and interoperability of metadata as widely as possible and cover the needs of a wide range of communities.

For further information, see the Workshop's Web site at:
http://hosted.ukoln.ac.uk/ec/metadata-1997/

The report can be downloaded from:
http://www2.echo.lu/libraries/en/metadata.html

For more information on the Libraries sector of the Telematics Application Programme, see:
http://www2.echo.lu/libraries/en/libraries.html

1. Objectives and context

This document is the report of the first Workshop on Metadata, held in Luxembourg on 1 and 2 December 1997.

DGXIII/E4, the Electronic publishing and libraries unit, is organising a series of workshops on the issue of metadata. Intended participation is from libraries sector projects within the Telematics Applications Programme and from projects in other TAP sectors and other programmes, both EU and national.

Objective: to establish a platform for co-ordination between projects concerned with metadata in a broad sense.

Under the current Framework Programme for RTD there are a number of projects concerned with metadata as such or with descriptions and descriptors of electronic documents. These projects will come across the same issues and problems and will benefit from concertation, as this will allow them to compare their concepts and approaches with others.

Objective: to make a wider European community aware of developments in the standards arena and stimulate feedback from the projects to the standards.

Developments in metadata in the Internet, specifically in Dublin Core, are moving fast. Some European organisations invest in participating in the Dublin Core workshops but not all have easy access to this activity. By inviting Dublin Core workshop participants to present the developments in the proposed workshops, a wider European audience can be informed on this subject. At the same time, models and experiences from the projects can be fed back into the standards arena.

The intended practical result of the workshop was the identification of key issues to be addressed and discussed on an on-going basis, thereby giving the projects a handrail in their research and specification phase.

This first workshop of 1½ days contained a tutorial, project presentations, breakout sessions discussing various aspects and resulted in conclusions and recommendations for further work as presented in sections 9 and 10 of this report. The programme is appended as Appendix 1. The participants list is given in Appendix 2.

A second workshop of one day, scheduled for mid-1998, will contain an update from the standards and reports from projects leading into discussion on solution scenarios.

2. Scope

Definition of metadata

The best definition of metadata is also very generic: metadata is data about data. In this definition the nature of the data being described by the metadata is left unspecified. This data can be collections or individual instances of objects or documents, Internet resources but also real world objects. By this definition, metadata also includes the cataloguing records that libraries have been producing for a long time.

In a more narrow sense, however, metadata commonly refers to information available on the Internet. This is also the environment where the necessity of good description mechanisms was becoming the most pressing: on the Internet and most notably the World Wide Web, it is relatively easy to publish information but it is becoming more and more difficult to find **relevant** information. Also, it is also very difficult to protect or control the use of information.

The metadata community

Many communities have a need for metadata of one type or another, including amongst others libraries, archives, museums, publishers, governments, information brokers. The interest in this workshop is an indication of the variety of interested parties, with representation from projects under various sectors of the Telematics Applications Programme (Libraries, Information Engineering, Language Engineering, Research).

Various sets of metadata

There are different requirements for metadata such as: discovery of information, protection of rights, classification of content, management of resources. Therefore, various communities might need different sets of metadata.

Currently, one of the most advanced sets is the one developed under the name of Dublin Core. Other sets are being developed such as that in PICS (Platform for Internet Content Selection).

This workshop had its focus on the Dublin Core metadata set, which is especially concerned with resource discovery of document-like objects on the Internet. It should be borne in mind that Dublin Core does not cover all possible requirements for metadata, but serves as a good and well-developed example, as well as being the one that is most relevant for the projects that were represented in this workshop.

3. The Tutorial

The tutorial was given by the UK Office for Library and Information Networking. Copies of the tutorial slides are included as Appendix 3.

The tutorial started with an introduction to the metadata landscape, describing:

- the variety of scenarios in which metadata appears (curatorial traditions, network resource discovery and network information management);
- the variety of operations supported by metadata (discovery, location, selection, acquisition, documentation, preservation, management, etc.); and
- the variety of models for metadata creation (author/creator, repository manager, third party).

Three bands of metadata formats were described, namely:

- the proprietary formats used by the big Web search engines;
- simple structured generic formats (Dublin Core and IAFA templates for example); and
- more complex, domain specific formats like MARC EAD and TEI headers.

Based on this overview, the tutorial described the Dublin Core in some detail, describing its positioning as a simple set of elements for resource discovery, set in an explicitly cross-sectoral environment. Some examples of Dublin Core were shown including a brief demonstration of UKOLN's DC-dot, a Web based tool for creating Dublin Core embedded into HTML META tags.

There are three models to manage metadata across a Web-site. These models are:

- the simple use of HTML authoring tools;
- the use of Web-site management tools; and
- the use of customised tools for holding metadata separately from the HTML Web pages and embedding it on the fly as the pages are served.

The third model, as implemented on the UKOLN Web server, was described in some detail.

A short history of Dublin Core followed, outlining the most significant steps on the way from the first meeting in Dublin, Ohio, USA (March 1995) to the fifth in Helsinki, Finland, in October 1997. Canberra Qualifiers, the way in which the meaning of the 15 Dublin Core elements can be refined, were described and a list of the current RFCs in preparation by the Dublin Core presented. The scope of Dublin Core implementation was briefly described (30 projects in 10 countries across a range of disciplines). Use of Dublin Core outside these projects is currently thought to be minimal.

The tutorial then went on to discuss interoperability issues, including protocol issues (WHOIS++, Z39.50), attribute naming and result sets. The use of Dublin Core to provide a consistent set of attribute names for searching and a consistent set of attributes for result sets was discussed. Some thoughts about registries and

some pointers to available software tools for converting between metadata formats were also given.

The process of harvesting, using a Web robot to gather information about Web based resources, and indexing was discussed. Again, some pointers to available tools were given (see appendix 3), including discussion of NWI (the Nordic Web Index) and other products.

The tutorial ended with a brief critique of Dublin Core and a look to the future. In particular, recent developments with RDF, the Resource Description Framework, and its relationship to the Dublin Core were described.

4. Presentations of current projects

Three presentations about current projects took place at the Workshop. The presentation slides are included in Appendix 4.

Hyperborea: Nordic Metadata and other Nordic Dublin Core initiatives **by Juha Hakala, Helsinki University Library, Finland**

This presentation covered the Nordic Metadata project and Nordic initiatives which utilise metadata creation, harvesting and indexing tools developed in the Nordic Metadata project.

Nordic Metadata is funded by NORDINFO. Its tool set includes a metadata template and user guides to it, a metadata-aware version of the Nordic Web Index (which is on its way to becoming a European Web Index) and DC->MARC converter. These tools are relatively mature, but development work is still continuing. All of the main aims have been fulfilled, but some minor things (like making the DC->MARC converter table-driven) cannot be accomplished in this project.

The Danish INternet DOCument REGistration (INDOREG) project is managed by the Danish Library Centre. The project intends to clarify some fundamental questions concerning bibliographic control with Danish Internet Documents. It is focused on the following areas in particular: inclusion criteria, registration method and metadata, PURL and publication of registrations (for more information, see the project report at http://purl.dk/rapport/html.uk/).

Another relevant Danish initiative is Netpublikationer, in which Dublin Core-based metadata is added into State official publications published in the Web.

Sweden, Norway and Finland do not yet have national initiatives like INDOREG, but there are strong local or regional projects. These have in some cases been born around Subject Based Information Gateways (one example is NOVA, Nordic SBIG for agricultural sciences and forestry), while other initiatives aim at enriching

research reports and dissertations published in the Web with Dublin Core metadata.

The major conclusion from these projects is that the creation and maintenance of high-quality metadata require a professional approach providing training for authors and other metadata creators, user guides and tools for automatic generation or computer-aided creation of metadata. As the speaker put it: "metadata as a 'hobby' does not work too well".

Furthermore, in the projects mentioned a very well developed technical infrastructure is available but the volume of metadata is still relatively small.

Metadata in CIMI and Aquarelle **by Mike Stapleton, SSL, UK**

This presentation covered the approach taken by CIMI and Aquarelle in the area of metadata, and specifically in the access to metadata already existing in the form of descriptive information of objects and archives in the field of cultural heritage.

The Aquarelle and CIMI/CHIO projects both seek to facilitate access to information in the heritage domain. Project CHIO defined a Z39.50 profile for information access between museums and carried out interoperability tests; Aquarelle aims at providing a service offering a single point of access to heritage information. Both are based on the use of gateways to existing data, dynamically creating "virtual metadata records" rather than collecting new metadata data. This difference should not be visible to the user but may be important to the data service provider.

Aquarelle and Project CHIO use similar but not identical data models as defined by their Z39.50 profiles. This presentation addressed issues encountered when defining the profiles and in particular the early reactions from the user and implementing organisations involved with the projects.

Major points identified are the problems in mapping existing descriptive sets to general metadata schemes. Metadata sets are classified in relation to the type of use. A set with only one element is useful for general searches, as currently demonstrated by the general search engines; full scale collection management systems will use sets of more than 200 elements to provide specific services to researchers with a detailed knowledge of the data set. The author argued that a general metadata set such as Dublin Core (15 elements) is most useful for researchers with a reasonable knowledge of the subject domain.

Towards interoperability: enabling resource discovery across the arts and humanities **by Paul Miller, Archaeology Data Service, UK**

This presentation reported on the results of a workshop series jointly conceived by the UK's Arts & Humanities Data Service (AHDS) and Office for Library and Information Networking (UKOLN), in which users' requirements for resource

discovery metadata were evaluated and compared with the developing Dublin Core initiative. The lessons learned by AHDS in actually implementing these findings within a working spatially, conceptually and disciplinarily distributed system were also addressed.

Much of the current work with 'metadata' is geared towards servicing the needs of experts WITHIN individual disciplinary communities. The problems faced in such work are multiplied when attempts are made to encompass the requirements of expert and non-expert across a range of loosely related disciplines, such as those comprising the arts and humanities.

Major conclusions are that metadata standards serve to make the access points to heterogeneous collections described using various descriptive practices comparable so the collections can be presented to the users through a single gateway, in this case to Arts & Humanities data.

Although the Archaeology Data Service (ADS) which the author is developing is using Dublin Core and demonstrates the validity of the approach, it is identified that the current state of Dublin Core is unsatisfactory because many elements are either confusing or insufficiently stable. This required the ADS project to define its own interpretations of the elements.

On the issue of gathering metadata from different sources, it was suggested that Z39.50 would be the solution for access because of the possibilities to express many data types and the experience in implementation.

Key conclusions from the project presentations

Summarising the issues addressed and the problems identified in the project presentations, the following points can be listed:

- metadata creation and maintenance requires a professional approach with appropriate organisational and technical basis: training, guidelines, tools
- the success of metadata initiatives is dependent on the availability of a critical mass of data
- integration of existing databases of descriptive information is important; the use of Z39.50 is considered to be an important tool to achieve this

5. General introduction to breakout sessions

In discussing metadata issues, three general areas can be identified where metadata is relevant.

First of all, metadata needs to be created or generated. This can involve manual intervention where authors, cataloguers or subject specialists assign descriptive

information to documents. It can also be based on automatic generation tools, for example based on document mark-up.

Secondly, metadata that has been created can be used for indexing and disclosure of information. This is the realm of organisations utilising harvesting tools and providing public access to the indexes.

Finally, the metadata will be used for searching and retrieval of resources.

According to the three areas of interest above, three breakout sessions were organised during the second day of the workshop.

Prior to the workshop, participants had been asked to indicate their interest in one of the three sessions and provide suggestions for issues to be covered. The list of issues identified for each of the three sessions is given in the reports of the individual sessions following.

6. Breakout session: Metadata creation and maintenance

Topics identified were:

- Management issues (models, management tools, embedded/separate metadata, hierarchical grouping, long-term availability, quality)
- Creation issues (automatic/manual creation, editing tools, guidelines for creators, import of metadata from publishers in SGML)
- Registries (use of schemes, controlled vocabularies, responsibilities for definition and maintenance)

The group started by trying to identify why participants were interested in creating metadata and what they wanted to describe. Reasons given included providing timely access to resources (in a news/alerting context), providing validation of resources (in a Health and Safety context) and providing a searchable index of resources (in an Engineering context). The group split into two sections to further identify issues - one group concentrating on the embedding of Dublin Core into HTML Web pages, the other on using Dublin Core as a way of unifying access to different collections of resources.

Discussions focused on the tools available for creating metadata and whether those tools easily fit into people's workflows, on the problems associated with the lack of stability of the Dublin Core itself, on the need for registries for various aspects of the Dublin Core and on the relationship between Dublin Core and Z39.50. Lack of significant deployment of Dublin Core and of guidelines for use was also discussed.

The group recognised the fact that the take-up of Dublin Core is slow and identified a number of inhibitors:

- lack of stability in Dublin Core

- lack of definitive/authoritative information
- lack of tools for creation/ harvest
- lack of deployment
- lack of protocol support
- lack of interchange formats
- lack of registries
- lack of guidelines
- lack of cross–sectoral representation

Solutions that were indicated were:

- Managed development path for DC
- Creation of registry process
- 'Partner–Find' programme
- Guidelines for dissemination and use.

In its conclusions, this breakout session recommended that the European Commission establishes contact with the Dublin Core community to request a clear statement of the current status and development plans of Dublin Core, and the timescales for a stable version. Furthermore, the Dublin Core community should provide adequate structures for organisations to join and participate in the discussions.

The European Commission was further asked to provide a statement of business benefits and constraints related to metadata creation and use within an EU context and to disseminate the resulting information widely across Commission-sponsored projects.

Finally, it was suggested that the Commission could promote the use of Dublin Core for dissemination of information about EC-sponsored projects.

7. Breakout session: Harvesting

Topics identified were:

- Process/execution and management issues (heterogeneous sets, maintenance of indexes, quality of links, long-term availability)
- Result issues (subject approach, multilinguality, minimalist/structuralist indexing, use of schemes)

The group used the categories of topics above to identify the major problem areas and discuss possible solutions.

As a first step a definition of harvesting was agreed saying that harvesting is basically the reverse operation of metadata creation, in the way that a harvesting

tool needs to disassemble the metadata into separate indexing categories, and that the primary goal of harvesting is to make the metadata useful for searching.

One major problem area for harvesting is the quality of the source metadata. There are two areas of interest:

- Building a general index across various subjects and domains
- Building subject- or domain-specific indexes.

It was identified that general indexes will always have a limited quality because of the diversity of the source material. Tools built into mainstream software packages such as text processors might increase quality.

In the area of subject- or domain-specific indexes, quality can be ensured by agreements across the interest community, for example through a maintenance agency for controlled vocabulary, specific tools and legal regulations (for example for government information for a particular country). The participants concluded that the establishment of such domain-specific agreements and the use of agreed vocabularies should be encouraged.

Regarding the level of indexing that would be necessary to help searchers to discover resources, there is a need for collection metadata as well as item metadata. A special category of metadata is metadata about resources that are not easily accessible because they are 'hidden' behind CGI (Common Gateway Interface) or other scripts or are accessible through special-purpose protocols. This metadata should be published in a way that the harvesters can easily get to use it.

The harvesting group also identified the fact that from a practical point of view, harvesters will only start using metadata when there is enough available.

The issues related to the harvesting of metadata for multi-lingual resources was not covered in the time allocated to the breakout session although the participants felt this was an important issue to be considered further.

8. Breakout session: Retrieval

This was a diverse group, coming from different sectors (libraries, museums, technologists) and having different interests in the types of data and resource to be retrieved (images, cultural resources, bibliographic services and data).

The topics that had been identified in advance and presented to the group as a baseline for discussion were:

- Precision vs. recall (full-text search versus metadata searching, use of sub-elements, use of schemes)
- Cross-domain issues (searching different sets of metadata, relation to Z39.50)
- Multilingual searching (user interface issue or harvesting issue?)

- Searching for collections of documents (levels of metadata)
- User guides/explain information (instructions for use, explanation of results)

In the event, most of the discussion revolved around various aspects of cross-domain searching, including 'levels' of metadata, and on user and usability issues. This did not mean that the other issues were seen as unimportant: multilingualism, in particular, was recognised by all as a critical topic for the future.

Early discussion focused on the need for and relevance of metadata, particularly as defined in the Dublin Core. A number of points emerged:

- concern for the potential for conflicts between existing schemes and solutions for interoperability in retrieval and the DC solution;
- recognition of the complexity of the domain;
- the present experience and application of DC are heavily rooted in the HTML environment (where a syntax for it exists) and this may be a limitation.

Overall, the group agreed that a common approach to metadata was necessary, especially to support cross-domain searching, across different types of objects and descriptive practices. However, in that context, it may be that even a simpler or 'shallower' set of descriptive elements is needed to answer simply the questions who, what, where and how. The model emerging in the discussions was one of essential metadata at a high level of extrapolation with drill-down features to richer and more specific schemes. On the specific issues of Z39.50, the existing Bib-1 attribute set and the relationships to DC, opinions were divided about whether DC should be accommodated within existing Bib-1 mapping or whether a separate DC attribute set should be established.

The second main focus to emerge in the discussion was usability, with the main points being:

- lack of information and knowledge of user behaviour in searching;
- a corresponding lack of understanding and knowledge of user expectations.

As a result, there is an identified need for systems to explain the scope of searches, the context for search results, and the nature of the objects and resources retrieved. More information and understanding is needed of the different layers in the information model and on their impact and use in the retrieval process.

A number of other topics were also identified as important for effective retrieval and delivery systems and which need to be addressed through the provision of metadata, though they fell outside the direct remit of this breakout session. They included the need for stable, authoritative and unique identifiers for objects; and the problems and information associated with rights handling and conditions of access.

9. Conclusions

The workshop participants recognised that metadata standards are necessary, in addition to sector- or subject-specific description mechanisms, to ensure interoperability in resource discovery on the Internet.

The workshop concentrated on the emerging and currently best-developed metadata format known as Dublin Core. It is however recognised that it concentrates fully on resource discovery and does not cover other requirements, e.g. for resource management or access restrictions.

The workshop, although recognising the usefulness of Dublin Core as a starting point in metadata descriptive standards, brought forward a number of concerns regarding the current state and the further development of Dublin Core:

- There is currently no formal responsibility for the maintenance of Dublin Core: development takes place in a very informal group of experts which meets once or twice per year in what is known as the Dublin Core Workshop Series.
- The current technical state of Dublin Core is unstable: during the meetings of the Dublin Core group, changes are being made to the format and there is no convergence to a stable version.
- The use of the current Dublin Core metadata format is not supported by the existence of guidelines: some of the philosophy and terminology of Dublin Core is not obvious to the uninitiated user which could lead to different interpretations adversely affecting interoperability.

It was also identified that the current take-up of Dublin Core is slow and that there is a lack of critical mass. This seems to be a classical chicken-and-egg situation: authors and publishers do not invest in providing Dublin Core metadata if the harvesters do not utilise it, and harvesters do not collect Dublin Core and use it for selective indexing if there is not enough data available. If this situation can not be changed, Dublin Core might not turn into reality.

10. Recommendations for further work

The workshop identified a number of actions that could be taken to promote the use of Dublin Core, including the following:

1. There needs to be clarity about the version control and the maintenance of Dublin Core. The Dublin Core group, addressed through the mailing list META2, will be asked to give a clear statement about this.

2. Further pilot projects should be started to further develop experience, test out the issues and help realise a critical mass of Dublin Core metadata. The European Commission and national bodies like National Libraries might have a role to play by encouraging the provision of Dublin Core metadata in documents, e.g. in project deliverables and electronic documents in the national deposit.

3. The interest and requirements existing in Europe warrant the establishment of a European group of implementers discussing the practical issues of implementing metadata in general and Dublin Core in particular. The Luxembourg Workshops, such as this December 1997 one, and a second one scheduled for mid-1998, could develop into a regular series.

4. The liaison with other groups concerned with metadata, such as the CEN/ISSS working group on Metadata for Multimedia Information (MMI), should be established to ensure applicability and interoperability of metadata as widely as possible and cover the needs of a wide range of communities.

11. List of acronyms and references

ADS	Archaeology Data Service: http://ads.ahds.ac.uk/ahds/
AHDS	Arts and Humanities Data Service: http://ahds.ac.uk/
Aquarelle	EC-sponsored project building an Information Network on Cultural Heritage for curators, urban and regional planners, publishers and researchers: http://aqua.inria.fr/
BIBLINK	Telematics for Libraries project aiming to establish a relationship between national bibliographic agencies and publishers of electronic material: http://hosted.ukoln.ac.uk/biblink/
CEN/ISSS	European Committee for Standardisation - Information Society Standardisation System: http://www.cenorm.be/isss/default.htm
CGI	Common Gateway Interface, a convention for external gateway programs to interface with Web server software
CHIO	CIMI project Cultural Heritage Information Online: http://www.cimi.org/projects/chio.html
CIMI	Consortium for the Computer Interchange of Museum Information: http://www.cimi.org/
DC	Acronym for Dublin Core
Desire	Telematics for Research project addressing the needs of research users in the context of a European information network based on the World Wide Web (WWW): http://www.surfnet.nl/surfnet/projects/desire/
DG XIII	Directorate General XIII of the European Commission: http://europa.eu.int/en/comm/dg13/13home.htm. See also: http://www2.echo.lu/home.html
Dublin Core	Dublin Core is a 15-element metadata element set intended to facilitate discovery of electronic resources: http://purl.oclc.org/metadata/dublin_core/
EAD	Encoded Archival Description: http://www.loc.gov/rr/ead/eadhome.html
EC	European Commission: http://europa.eu.int/
HTML	Hypertext Mark-up Language, the document format for WWW documents
IAFA	Indexing information specified by the working group of the IETF concerned with Internet Anonymous FTP Archives that can be used to describe the contents and services provided by such archives: http://info.webcrawler.com/mak/projects/iafa/iafa.txt
INDOREG	The Danish INternet DOCument REGistration (INDOREG) project: http://www.purl.dk/rapport/html.uk/
ISO	International Organisation for Standardisation: http://www.iso.ch/
MARC	Standard format for the representation and communication of bibliographic and related information in machine-readable

	form, ISO standard 2709
MMI	CEN/ISSS Workshop on Metadata for Multimedia Information: http://www.cenorm.be/isss/Workshop/MMI/welcome.html
NISO	U.S. National Information Standards Organization: http://www.niso.org
Nordic Metadata	Scandinavian co-operation project creating basic elements of a metadata production and utilisation system: http://renki.helsinki.fi/meta/
NORDINFO	The Nordic Council for Scientific Information
NOVA	Nordic SBIG for agricultural sciences and forestry
NWI	Nordic Web Index: http://nwi.bok.hi.is/; http://nwi.dtv.dk/; http://nwi.bibsys.no/; http://nwi.funet.fi/; http://nwi.ub2.lu.se/
PICS	Platform for Internet Content Selection, an infrastructure for associating labels (metadata) with Internet content: http://www.w3.org/PICS/
RDF	Resource Description Framework, a specification currently under development, designed to provide an infrastructure to support metadata across many web-based activities: http://www.w3.org/RDF/
RTD	Research & Technological Development
SBIG	Subject Based Information Gateways
SGML	Standard Generalized Markup Language, ISO standard 8879
TAP	The Telematics Applications Programme, one of the European Commission's research programmes, aimed at stimulating RTD on applications of information and/or communications technologies in areas of general interest: http://www2.echo.lu/telematics/telehome2.html
TEI	Text Encoding Initiative, an international project to develop guidelines for the preparation and interchange of electronic texts for scholarly research: http://www.uic.edu/orgs/tei/index.html
Telematics for Libraries	The Libraries sector of the Telematics Application Programme: http://www2.echo.lu/libraries/en/libraries.html
UKOLN	The UK Office for Library and Information Networking: http://www.ukoln.ac.uk/
Z39.50	NISO standard specifying an Open Systems Interconnection application protocol for Information Retrieval, identical with ISO standard 23950: http://lcweb.loc.gov/z3950/agency/document.html

Appendix 1. Programme

1 December (14:00-18:00)

14:00-14:15 Welcome - Ariane Iljon, Head of Unit DGXIII/E-4

14:15-14:30 Introduction on scope and objectives - Makx Dekkers

14:30-18:00 Metadata tutorial, organised by the UK Office for Library Networking

- *Introduction*
- *The metadata landscape*
- *Review of Dublin Core*
- *Data creation and harvesting*
- *Interoperability*
- *Future*

2 December (09:00-16:00)

9:00-10:00 Project presentations

- *The Nordic Metadata Project* - Juha Hakala, Helsinki University Library, Finland
- *Metadata in CIMI and Aquarelle* - Mike Stapleton, SSL, UK
- *Towards interoperability: enabling resource discovery across the arts and humanities* - Paul Miller, Archaeology Data Service, UK

10:00-13:00 Breakout sessions with assigned rapporteurs on the following subjects:

- *Metadata creation and maintenance*. Chair: Lorcan Dempsey
- *Harvesting*. Chair: Sigfrid Lundberg
- *Retrieval*. Chair: Gordon Pedersen

13:00-14:00 Lunch

14:00-15:00 Breakout session reports by rapporteurs

15:00-16:00 Discussion, resolutions chaired by Ariane Iljon

Appendix 2. List of participants

METADATA WORKSHOP Luxembourg
1-2 December 1997

List of Participants

Mr. Philippe AIGRAIN, CEC - DG III/F6	Belgium
Mr. Stephen P. BAGLEY, The British Library	UK
Ms. Stefania BIAGIONI, IEI - CNR	Italy
Mr. Ian BLOOR, De Montfort University	UK
Mr. Peter BURNHILL, University of Edinburgh Computing Centre	UK
Mr. Steven CARR, SCRAN	UK
Mrs. Donatella CASTELLI, IEI - CNR	Italy
Mrs. Robina CLAYPHAN, The British Library	UK
Ms. Patrizia COTONESCHI, Università degli Studi di Firenze	Italy
Mr. Makx DEKKERS, Coopers & Lybrand	Luxembourg
Mr. Lorcan DEMPSEY, UKOLN - University of Bath	UK
Mr. Werner DEUTSCH, Austrian Academy of Sound Sciences	Austria
Mr. Jacques DUCLOY, INRIA Lorraine - CRIN CNRS	France
Mr. Robbert FISHER, Coopers & Lybrand	Luxembourg
Mr. Esben FJORD NIELSEN, Aarhus Kommunes Biblioteker	Denmark
Mr. Kevin P. GARDINER, CEC - DG XIII/E4	Luxembourg
Mr. Stefan GRADMANN, Pica	Netherlands
Mr. Juha HAKALA, Helsinki University Library	Finland
Mr. Paul HARVEY, Fretwell - Downing Data Systems Ltd.	UK
Mr. David HITCHCOCK, KET	Luxembourg
Ms. Ariane ILJON, CEC - DG XIII/E4, Head of Unit	Luxembourg
Mr. Poul Henrik JØ RGENSEN, Portia System I/S	Denmark
Mrs. Ulrike JUNGER, Niedersächsische Staats-und Universitätsbibliothek	Germany
Mr. Peter KESTELOOT, WTCM	Belgium
Mr. LECHAUDEL, INIST	France
Mr. Dominique LERINCKX, ULB - Library	Belgium
Mr. Hans LIEGMANN, Die Deutsche Bibliothek	Germany
Mr. Sigfried LUNDBERG, Lund University Library - Netlab	Sweden
Mr. Luigi MADELLA, Italia Online SpA	Italy
Ms. Cristina MAGLIANO, Istituto Centrale Catalogo Unico	Italy
Mrs. Pat MANSON, CEC - DG XIII/E4	Luxembourg
Mr. Franco MASTRODDI, CEC - DG XIII/E	Luxembourg
Mr. Paul MILLER, University of York - Archeology Data Service	UK
Mr. K.P. MILLER, University of Essex - ESRC Data Archive	UK
Mr. Charles-Henri NYNS, Université Catholique de Louvain	Belgium
Mr. Gordon PEDERSEN, Fischer & Lorenz	Luxembourg
Mr. Ian PIGOTT, CEC - DG XIII/E4	Luxembourg
Mr. Alain PIROTTE, Université de Louvain	Belgium
Mr. Thomas PLACE, Tilburg University	Netherlands
Mr. A. POWELL, UKOLN - University of Bath	UK
Mrs. Maria Angeles ROJAS, CEC - DG XIII/C3	Belgium
Mrs. Mary ROWLATT, Chelmsford Library	UK
Mr. ROZENFELD, ISSN International Centre	France
Mrs. Monika SEGBERT, The British Council	Germany
Mr. SIMONETTA, INIST	France
Mr. Peter SMITH, LASER	UK

Mr. Pierre Paul SONDAG, CEC - DG XIII/E5	Luxembourg
Ms. Janet STAFFORD, University of Sunderland	UK
Mr. Mike STAPLETON, System Simulation Ltd.	UK
Mr. Hans VAN DER NEUT, Rijksarchiefdienst	Netherlands
Mr. Paul VERMEIREN, CEC - DG X	Belgium
Mr. Per VESTBOSTAD, University of Bergen	Norway
Mrs. Jutta WEBER, Staatsbibliothek zu Berlin	Germany
Mr. WEETS, CEC - DG XIII/C3	Belgium
Mr. Justin WICKENS, CEC - EUROSTAT	Luxembourg

Appendix 3: Presentation slides of Tutorial

Appendix 4: Presentation slides of project presentations