The Digital Object Identifier System, the Dublin Core Community and the INDECS Project: Converging metadata initiatives for describing information resources as intellectual property

JEAN MAREE
ASSISTANT- DIRECTOR and ACQUISITIONS LIBRARIAN
DEPARTMENT OF LIBRARY SERVICES
UNIVERSITY OF SOUTH AFRICA
30 MAY 1999

ABSTRACT: This paper describes the Digital Object Identifier (DOI) system initiative to add metadata to the original DOI resource identifier concept. This converges with the acknowledgement by the Dublin Core (DC) community that resource description metadata should also embrace other aspects of resources, such as rights and copyright. The metadata models (Dublin Core, INDECS) and tools (FRBR, RDF, XML) which influenced this convergence are also described, as well as the role of the INDECS project in brokering agreement on an interoperable metadata model for describing information resources as intellectual property for digital transaction processing.

"Libraries want to share content; publishers want to sell it. Museums strive to preserve culture, and artists to create it. Musicians compose and perform, but must license and collect. Users want access, regardless of where of how content is held. What all of these stakeholders share is the need to identify content and its owner, to agree on the terms and conditions of its use and reuse, and to be able to share this information in reliable ways to make it easier to find."

David Bearman et al.: A common model to support interoperable metadata. *D-Lib Magazine*, January, 1999.

CONTENTS:

1: INTRODUCTION

1.1: Problem statement:

Creators, publishers and other rightsholders now find themselves trading electronically and across geographical borders with simultaneous physical and digital versions of their intellectual property. This complicates an already byzantine landscape of complex rights ownership and administration.They realise that they need a common data structure to manage their rights, interests and services in networked environments, so that content integrated with automated services may improve many aspects of a highly fragmented industry.

1.2: Some problems raised by online intellectual property

Firstly, legal copyright protection becomes difficult to apply to digital resources, since the fundamental criteria to date have been original intellectual effort and fixation in tangible forms such as medium of expression, minimum length or quantity (Lyons 1997).

On the Internet, digital information, though initially fixed as machine-readable data (Oppenheim 1998:12), becomes infinitely malleable, since it can be interactively accessed, streamed and intermingled from different sources, and dynamically generated and processed on the fly to create new works and versions of works. The American Bar Association currently uses a working definition of digital works as literary works which are 'capable of behaviour', 'performed by being mapped into waveforms', and where usage may be defined as "access to perform stated operations upon a sequence of bits" (Lyons 1997).

As with print or physical creations, however, owners or creators of digital content are still entitled to the 'moral right' of correct attribution of their creations, to protection against 'derogatory treatment such as cropping or amending of their works, and to compensation such as royalties.

Content providers struggle with product definitions to obtain copyright protection: for instance, online databases have only recently been recognised in law by a European Union Directive as a special literary type, with certain conditions for receiving a 15-year overall database right protection, regardless of the copyright that may or may not reside in its individual components (Oppenheim 1998: 31). Multimedia are not necessarily simple compilations, but could be seen as configurable computer programmes or databases.

Electronic distribution of information products also poses problems, such as unauthorised

access and misuse of material, e.g. copying and posting of copyrighted material to open networks, or the sometimes illegal distribution of tables of content in SDI services.

Publishers are also concerned with the management of versions (physical, digital, mirrored, cached) of original works, and with the functional granularity issue, whereby parts of works may be traded separately for reuse in other works such as anthologies, coursepacks, electronic reserves and multimedia productions.

Users, both private and in public, academic, and corporate research libraries, want fair use of material, networked widely and conveniently by their parent organisations; they want to view, listen to, browse, copy to disk, transmit, print, download, or cut and paste material into works which they might create in turn (Alrashid et al 1998). Distance educators may want to show videoclips or play music recordings to remote students not located in formal classrooms; these students in turn may download or save these works to disk. Users may want privacy of information usage and may object to customer tracking and identification (Davis 1997). They are not always copyright aware, and may often transgress the boundaries of permissibility.

Returning to the law, which most would suppose to require precision in legal transactions, some legal experts have warned against the dangers of mechanical enforcement of terms and conditions for information access; according to them, room for 'ambiguity' or at least for interpretation in definitions is also required. This can be illustrated by reference to the concept of fair dealing (UK), fair use (USA), or "private copying" (Europe) which has no precise definition, and is often determined by interpretation in context. If access to on-line material is only possible through mechanical enforcement, then fair use becomes impossible (Davis 1997). Where fair use has been abused, the crime is infringement, and the penalty can be damages in a civil court action; flagrant abuse showing a deliberate intent to harm the legitimate interests of the copyright holder can lead to a criminal case (Oppenheim 1998: 20).

All of this is increasingly difficult to administrate in online commerce. Publishers are therefore interested in joining forces to find a common model of product identification and description that could enable as much automated rights and permissions trading and protection as possible by means of intelligently executed business decision models, 'trees' or templates.

Slide 1 shows some of the elements that need to be identified in automated intellectual property transactions:

SLIDE 1: ELEMENTS WHICH FEATURE IN INTELLECTUAL TRANSACTIONS

(from BIC/Editeur/NISO rights metadata working party, interim report 6/3/98)

OBJECT USE USER

work(e.g.,encyclopaedia) personal licensee:

part of work (entry in enc) corporate individual:

version (trans/edition) commercial organisation

manifestation(print/dig) non-commercial (org. agent?)

component(chapter) in what form: end user

part of component(illustr) embedded? closed user group?

embedded object(table/slide) alone? open user group?

part of embedded object version of copy? roles?

manifestation of copy? status?

limitations: payment method?

territory

duration

exclusivity

1.2: **The identification of intellectual property**

The first principle of intellectual property management is the unambiguous and enduring identification of each item needing protection.There are two steps in this process:

1. Assign a name or identifier : A seminal paper in this regard was that of Kahn and Wilensky (1995) on distributed digital objects, which described a digital storage and access protocol in which online resources might feature as unique 'objects' whose identifiers are known as 'handles' and which are securely stored in global repositories.

Publishers have adapted this idea to suggest a global identifier registration repository, which could store their product identifiers and which could resolve each identifier into a secure access path to an online location where a digital object may be found as a structured package of often encrypted information, incorporating data about itself such as terms and conditions for access, business rules for rights negotiations, etc. This is the basis of the DOI system, which will be described later.

2. Assign key metadata: The context in which objects are declared unique still remains a human decision, subject to rules or laws (Gladney 1998). It is these formal specifications for the copyrightable uniqueness of intellectual property objects which are being explored by the metadata initiatives described below. Once there is agreement on these specifications, rights and permissions metadata can be added to the descriptive metadata of objects, and interposed between users and objects as a form of online intellectual property protection.

Any digital transaction must also identify users of objects; and must also identify authenticated non-repudiable transactions, whether commercial (payment, licensing, royalties, permissions) or 'cultural' (loosely defined as involving no payment). These aspects must also be described by means of accurate metadata.

2. THE DUBLIN CORE METADATA INITIATIVE

2.1: What is metadata ?

The Greek word "meta" means "that which lies beneath or beyond, or that which underlies something". A working definition of metadata is that it is "data about data". More helpful, perhaps, is Lynch's definition of "information that qualifies other information " (Lynch 1998).

Basic metadata about a resource can be inserted into the META tag of the HTML coding of a Web document. This tag has two description areas: the META description tag, where you can enter a brief description of your Web page; this could appear in a metadata-enabled search engine's results page. Similarly, by inserting meaningful keywords into the META keyword tag, you can also control the indexing of your Web page. Not all search engines support searching of the META tag, but it is imperative for future search engine design to improve the 'high recall, low precision' syndrome.

Resource description metadata is an expanded form of metadata, where the META tag is used repeatedly to enter considerable detail about a resource, thereby forming a label or a type of catalogue record, usually embedded in the coding of the resource itself or sometimes stored in a separate database.

2.2: Dublin Core resource description metadata

The dominant metadata set is Dublin Core metadata, which originally catered for simple resource description, but is now reaching popular status as a professional description standard within the library, archive, museum, and other communities.

DC description metadata is restricted to a set of 15 elements or tags, although there is a "structuralist" school which adds DC Qualifiers to each element. The 15 elements describe three aspects of a digital resource. As you may see, there is an attempt to describe the intellectual property aspect of a resource (elements 8, 9, 10 and 11):

SLIDE 2: DUBLIN CORE BASIC DESCRIPTION ELEMENTS

Content of resource:

1: Title or name given to resource by creator

2: Subject: simple keywords, or terms from a controlled vocabulary

3: Description: a textual description of the content

4: Source: the resource from which the current resource is derived

5: Language: language of the content; use language tags defined in RFC 1766

6: Relation: the identifier of another resource and its relationship/link to current resource

7: Coverage: temporal or spatial characteristics of the content, e.g. C19 South Africa

Intellectual property aspects of resource

8: Creator: person/organisation primarily responsible for intellectual content

9: Publisher: the body making resource available in its current form

10: Secondary contributor who has made significant contributions to the resource

11: Rights: a simple rights statement about the resource, or an identifier linking to information

The particular instance of the resource:

12: Date: a date associated with the creation or availability of the resource

13: Type: the category of the resource: a web page, article, etc

14: Format: the resource data format, e.g. MIME type; defines tools needed to operate on resource

15: Identifier: which uniquely identifies the resource; URL, ISBN, etc

Information communities often declare and construct metadata templates for consistent description, e.g. a template might require the following elements for describing an online journal article:

SLIDE 3: UNIVERSITY OF ARIZONA METADATA TEMPLATE: ONLINE JOURNAL ARTICLE

JOURNAL ARTICLE DESCRIPTIVE TAGS

<Title>.....</title>

<creator>.....</creator>

<Subject>.....</Subject>

<Abstract>.....</Abstract>

<Jrnltitle>.....</Jrnltitle>

<Jrnlvol>..... <Jrnlvol>

<Jrnlno>.....</Jrnlno>

<Jrnldate>..... .</Jrnldate>

<Pagefirst>.....</Pagefirst>

<Pagelast>.....</Pagelast>

<Pages>.....</Pages>

<Publisher>.....</Publisher>

<Contributor>.</Contributor>

<Type>.....</Type>

<Format>.....</Format>

<Identifier>.....</Identifier>

<Language>.....</Language>

<Location>.....</Location>

<Copyright>....</Copyright>

<URL>.....</URL>

A fully DC- described web- document's META tag looks like this in the HTML source page of Paul Miller's article in the digital periodical *Ariadne*:

SLIDE 4: DUBLIN CORE COMPLIANT DESCRIPTION

<!DOCTYPE HTML PUBLIC "-IETF//DTD HTML 2.0//EN">

<HTML>

<HEAD>

<TITLE>Metadata for the masses</TITLE>

[DUBLIN CORE METADATA ELEMENTS INSERTED INTO HTML META TAG]

<META NAME="package" CONTENT="(TYPE=begin) Dublin Core">

<META NAME=**"DC.title**" CONTENT="(TYPE=long) Metadata for the masses: what is it, how can it help me, and how can I use it?">

<LINK REL=SCHEMA.dc
HREF="http://purl.org/metadata/dublin_core_elements#title">

<META NAME=**"DC.title**" CONTENT="(TYPE=short) Metadata for the masses">

<LINK REL=SCHEMA.dc
HREF="http://purl.org/metadata/dublin_core_elements#title">

<META NAME=**"DC.subject**" CONTENT="(SCHEME=keyword) Dublin Core, Metadata, Warwick Framework, Resource Description, Resource Discovery">

<LINK REL=SCHEMA.dc
HREF="http://purl.org/metadata/dublin_core_elements#subject">

<META NAME=**"DC.author**" CONTENT="(**TYPE=name)** Paul Miller">

<LINK REL=SCHEMA.dc
HREF="http://purl.org/metadata/dublin_core_elements#author">

<META NAME="**DC.author**" CONTENT="**(TYPE=email)**
A.P.Miller@newcastle.ac.uk">

<LINK REL=SCHEMA.dc
HREF="http://purl.org/metadata/dublin_core_elements#author">

<META NAME=**"DC.author**" CONTENT="**(TYPE=postal**) University Computing Service

University of Newcastle Newcastle upon Tyne NE1 7RU UK">

<LINK REL=SCHEMA.dc
HREF="http://purl.org/metadata/dublin_core_elements#author">

<META NAME="**DC.author**" CONTENT="**(TYPE=phone**) +44 191 222 8212">

<LINK REL=SCHEMA.dc
HREF="http://purl.org/metadata/dublin_core_elements#author">

<META NAME="DC.author" CONTENT="(TYPE=fax) +44 191 222 8765">

```
<LINK REL=SCHEMA.dc
```

Communities using DC metadata are the German Digital Libraries Project, the E-lib projects in Britain, the USA Digital Libraries project, projects in Australia and the NordicWeb countries; the latter have also developed the first Dublin Core indexing "spider" which browses DC elements to compile its database. DC has been used for pre-print servers, ejournals, dissertations, subject gateways, national archives, art and museum servers and electronic text centres, and has been proposed as a standard to the IETF.

Whilst DC metadata may satisfy the resource description and discovery communities, the question is whether it is can describe resources to be accessed and traded as intellectual property. The current answer from the content owning and trading industries, or publishers, is No.

3. THE INTEROPERABILITY OF DATA FOR E-COMMERCE PROJECT

3.1: The INDECS stance on metadata

INDECS is a European Community INFO2000 project, which co-ordinates the work of many rights-holder groups, all concerned with eliminating trade barriers and fragmentation in the content industry as a whole. Some of these groups are: the DOI community, representing the book, journal and bibliographic industries, such as BIC/Editeur and their TIP Title Information Project for Bookdata and Whitakers; MUSE, an EC recording industry initiative; ISAN, a film industry initiative working on the International Standard Audio-visual Number; and the Common Information System (CIS), a copyright programme for licensing and royalty collection. INDECS also draws on other rights-related projects such as the graphic and performing communities.

These industries want common systems which enable "the description and rights ownership of digitally-traded creations" (Rust 1998a), because they must protect and trade rather than merely describe their creations in the online environment: Godfrey Rust, the INDECS spokesman, refers to "the raging fire of e-commerce, ready to burn up all before it" (Rust 1998b), and to the competitive advantage which arises from integrating content with services for direct interaction with information consumers.

The first task of INDECS is to design a common content identification model to use consistently and regardless of format type across all the content industries for their various creations such as books, journals, CD-ROMs, music compositions and recordings, television productions, electronic multimedia and films; the model must also allow for new online content 'hybrids' which may arise.

Once the model for identifying intellectual creations has been accepted, a standardised metadata vocabulary will be proposed. The project-leaders have all agreed that "the functional requirements of managing intellectual property rights include the ability to encode descriptive data at a very high level of precision..[whereas]..the description requirements for resource discovery are generally less precise" (Weibel 1999).

Rust argued strongly that DC metadata cannot accommodate intellectual property description requirements, warning that it would cause confusion about the exact version or part of the resource being traded. For instance, he objected to the treatment of the Creator, Contributor and Publisher as descriptive elements rather than as active rights-owning parties; and he pointed to the Source and Relationship aspects of DC which do not properly address the complexity of relationships between works and manifestations of those works, between works and parts of works, between works embedded in other works, e.g. a video clip used in a scientific article, and fully multimedia works where different rights may apply to each component.

The Dublin Core community has now undertaken to work on a Schema Harmonisation project

with INDECS, which may result in significant revisions of "DC1.0", having agreed with Rust that "there is no logic in having one core metadata set for discovery and one for rights management" (Rust 1998).

3.2: The INDECS metadata models for intellectual property description and transaction

INDECS has produced two models for testing against all types of intellectual creations, and will map them to corresponding legal definitions, after which they will produce a standardised data dictionary embracing all creation types. They will also shortly map their definitions to MARC and DC elements, and will produce XML expressions of their data elements in the new RDF framework.

Though not yet finalised, the contrast in precision between the INDECS metadata models for rights description and the Dublin Core model for resource description should be clear.

The first model is a general model, which outlines the Events related to the creation and trade of an item, which shows the interplay of 3 things: the objects of the transaction, the people or parties to the transaction, and the actual terms of the transaction.

SLIDE 5: BRIEF GENERAL MODEL OF INTELLECTUAL PROPERTY EVENTS: THE CREATION, ITS CREATORS AND THE TRANSACTION IN WHICH IT MIGHT FEATURE (each Event may be modelled in more detail, as shown in the subsequent slides)

Personal Personal Event Person/s event(birth/death)(nb70 yrs copyright

Creation-based Creative Event Event involving contribution to creation

Usage Event Event in which creation is used

Performance Event Creation which is an event

Transactions Agreement Event permitting use of creation

Offer Event making terms of agreement

Payment Event in which person pays money to another

The second model is a series of specific models for describing these Events in more precise detail. Notable here is the absence of genre or format categories, since the objective is to find common attributes for all creations. Looking closely again, the attributes are perhaps more production- orientated rather than bibliographic.

SLIDE 6: SPECIFIC MODEL: PERSONS ATTRIBUTES (creators, owners, rights-owners)

Attribute Attribute type Attribute Definition

Label Identifier Unique label given to person by designated authority

Name Text string by which person is known

Descriptor Description applied to person

Class Primary type Principle form of a person:natural,legal,corporate

Gender Sex of person, animal or apparent sex of quasi-person

Corporate Person legal status, apparent legal status of person

Quasi-person A form of quasi-person (pseudonyms,Teletubbies)

Animal Form of animal,etc (e.g. "Babe", "Lassie")

Context Principle activity associated (art,opera)

Extent Quantifiable attr of person/s: natural,legal,etc

Dimension Spatial characteristics (residence, company location)

Duration Temporal characteristic (born/died/company liquidate

Enumeration Enumeration of component (ensemble members)

Relation Relation betw entities (writers group,art-movement)

SLIDE 7: SPECIFIC MODEL: CREATIONS ATTRIBUTES ('production details')

Creation defined as "any thing or event in which intellectual property rights may exist"

Attribute Type Attribute Definition + examples

Label Text string identifies creation

Identifier Label allocated to creation by authority: ISBN,UPC

Title Name by which creation is known:title,brief title

Descriptor Text string describing creation: keyword,annotation

Class Elements of form,structure,content of creation

(New hybrid modes, formats, encodings, e.g. multimedia may arise)

Type Mode,form,content: audio recording,music work

Mode Sensory mode for perceiving creation:audio,visual

Format Structural manifestation: CD,paperback

Material Substance of manifestation: vinyl,paper,bits

Encoding How work is fixed in manifestation: print,MPEG2

Form Perform/work structural type:play,symphony,advert

Genre Style expressing performance/work:jazz,film noir

Language In which perform/work is expressed:Eng,braille,java

Subject Subject of work:chemistry,blues music

Context Purpose,audience,environm of creation:live studio

Continuity* Status of creation over time:series,sequel,interactive

Development Stage of creation evolution:draft,remix,reprint

Extent Measurement Elements of creation measurement

Dimension Spatial characteristic of creation:height, capacity

Duration Temporal characteristic of creation: duration

Enumeration Count components of creation:tracks,pages,verses

Relation Between creation + other entity:excerpt,version

Each attribute could be detailed in even greater depth to express all possible variations of that attribute, e.g. the **Continuity*** attribute of a Creation (slide 7 above) could be classified further as

SLIDE 8: SPECIFIC MODEL: CREATION CONTINUITY: finer detail

Attribute Type Example

Static Form/Content fixed Book, audio CD, photograph, Word file

Cumulative Content added over time Series, serial

Dynamic Form/content changes Webpage, streamed broadcast, stock prices

SLIDE 9: SPECIFIC MODEL: CREATION ROLES

(Role played by one creation in making of another creation, e.g. the film Shakespeare in Love was a manifestation of the pre-existing novel)

Role Definition

Original Creation not dependent on another/pre-existing creation

Compilation Creation made from comp. of 2 or more pre-existing creations

Constituent Constituent of new compilation

Version Creation made through transformation of pre-existing creation

Excerpt Creation made by extracting part of pre-existing creation

Replica Creation made by copying pre-existing creation

Source Creation from which a version,excerpt or replica is derived

Abstraction Work discerned from performance/ manifestation:symphony

Expression Performance/manifestation of pre-existing work

Root Pre-existing abstract work used in new performance/manifestation

Subject Creation made the subject of a new creation

Tool Creation as instrument in making another creation

SLIDE 10: SPECIFIC MODEL FOR CONTRIBUTOR VALUES

(persons who may possibly also have rights in a creation because of a contributing activity)

Role Definition Example

Originator makes orig creation author,composer

Modifier makes version arranger,translator,remixer

Compiler makes compilation anthologist

Extractor makes excerpt e.g., readings for television poetry program

Replicator makes replica printer, copier

Producer of final creation form publisher, studio producer

Director directs creation making choreographer, art director

Performer performs,interprets work orchestra, ensemble,guitarist

Recorder records creation event typist,cameraman,sound engineer

Operator operates equipment animator,special effects

Facilitator provide ancillary service researcher,financier,makeup artist

SLIDE 11: SPECIFIC MODEL: USER VALUES

(making other creations from existing creations)

User type Definition Example

Modifier/transformer Person using creation to make a version Arranger,adapt,remix

Compiler Person using 2 or more creations to compile Anthologist

Extractor Person using creation to make excerpt

Replicator Person using creation to make replica Printer,copier

Disseminator Person making creation available to others Publisher,distributor

Perceiver/Enduser Person who 'enjoys' creation in some mode
Reader,browser,listener

Keeper/retainer Person who retains possession of creation

Once the creation, its creators, its prospective user and his purpose have been identified, it should be possible to structure an automated transaction, e.g. purchase, license, permission to copy or translate. INDECS is mapping a high-level transaction model against 'real-world' deals for testing. Eventually it should be possible to present permissions applicants with a use/ purpose menu driven by a codified agreement template and customised with object-specific terms of availability, such as prices or permitted uses, e.g. educational, corporate research (for an example of a menu-driven rights transaction site, see http://www.copyrightdirect.com which is run by Yankee Book Peddler). The template design would obviously be user-friendly, but behind it will lie structured business data which may look like this.

SLIDE 12: GENERAL MODEL: TRANSACTIONS

AGREEMENT

Event Identifier= as assigned

Event Type= Agreement

Entity Type= Time= 19990409

Entity Type= Place= Identifier= www.johnsmith.com (publ website)

Entity Type= Person= Identifier= John Smith Publishers

Role= grantor

Entity Type= Person= Identifier= Bill Brown

Role= grantee

PERMITTED ACT AUTHORISED BY AGREEMENT

Event Identifier= as assigned

Event Type= Usage event

Entity Type= Time= as agreed

Entity Type= Place=Identifier= www.johnsmith.com (publ website)

Entity Type= Person= Bill Brown

Role= downloader= Purpose= private

Entity Type= Creation

Type= text file= identifier= title, DOI

CONSIDERATION (e.g., PAYMENT) FOR PERMITTED ACT

Event Identifier= as assigned

Event Type= Payment

Entity Type= Time= before (date)

Entity Type= Place= Identifier= John Smith Publ. bank account

Entity Type= Person= Identifier= Bill Brown

Role= Payer

Entity Type= Person= Identifier= John Smith Publ.

Role= Payee

Entity Type= Thing= Identifier= Money

Extent= $5.00

4. THE DIGITAL OBJECT IDENTIFIER SYSTEM

4.1: What is the Digital Object Identifier System?

The DOI system was developed at the request of the Association of American Publishers, with participants such as Wiley, Academic Press, Elsevier and Springer, all of whom are now represented by the International DOI Foundation (IDF).

It is a system whereby publishers may apply for permanent prefixes from a DOI registration agency , and then use these prefixes to construct unique DOI identifiers for their products; these identifiers must be deposited in the online DOI registry along with the URL addresses of the items. Clicking on a DOI or querying the DOI registry will enable resolution by the Corporation for National Research Initiatives (CNRI) Handle servers to the web-location of those products, where access, service or trade may take place. Publishers must undertake to update the DOI registry if the ownership or web-addresses of their products change. The Handle protocol always resolves to the latest address of the object as listed in the registry. Publishers may use their in-house production numbers, or may incorporate 'legacy' numbering systems, such as the ISBN, SICI or BICI into the DOI string if they wish.

Since the DOI is persistent, it is also preferred to the URL type of bibliographic citation, and it can enable more reliable navigation between resources, e.g. from abstract to full-text.

SLIDE 13: DIGITAL OBJECT IDENTIFIER

A typical DOI looks like this:

prefix **/** suffix

10. 1234**/** [ISBN]0-333-45678-x

i.e., a namespace identifier + namespace specific string [when seen as a URN or Uniform Resource name]

- prefix indicates the official DOI registration agency (10) [namespace identifier]

- and the prefix holder, publisher or copyright holder (1234) [namespace specific string]

- the suffix after the slash indicates the numbering scheme used by the publisher, i.e. [ISBN] - unique product number itself (0-333-45678-x) [namespace specific string]

The DOI registry translates the DOI number into the http protocol, and routes user to a URL location for the product/s, e.g.

http://www.macmillan.uk/mathchem99/smith

The INDECS project pointed out that the DOI would be more useful if structured into intellectual property descriptions of objects, where copyright and terms of access are attached to these objects. They also pointed out that resolution to multiple manifestations of Works (physical and digital) and their different terms of trade would be required. A DOI may now identify a printed item as well as its online version.

Whilst this has meant a delay in the implementation of the DOI system whilst a suitable metadata schema is agreed upon, the IDF is still confident that it will become the preferred identifier and access method for the intellectual content communities.

4.2 : DOI intellectual property identification metadata

The metadata model which the IDF now proposes to use is aligned with the INDECS models, although only a small core description of direct item attributes, called a "kernel", will be associated with the object when it is registered with the DOI agency, rather than a full INDECS Creation Event description. DOIs accompanied by metadata are known as Level 2 DOIs. For the sake of consistent metadata for each type of creation, broad genre classifications (journals, sound recordings, videos) will probably be used. This is a pragmatic decision, since different DOI agencies might best administrate the registration of different products.

Below is the proposed metadata kernel which should accompany any item when it is registered with a DOI agency:

SLIDE 14: DOI METADATA KERNEL

Element Definition Status Value/Qualification

DOI a DOI Mandatory DOI

DOI Genre Class of resources Mandatory DOI Genre Tables

defined by IDF

Identifier Unique resource ID Qualified String

Title Resource name Qualified String

Type Structural type of Mandatory Work/ Phys/ Dig/ Performance

resource

Origination Process by which Mandatory Orig/ Deriv/ Extr/ Repl/ Comp.

resource is made

Primary Agent Name or ID of Mandatory Genre Namespace ID/ Name

primary agent/s

Agent role/s Role/s played by Mandatory Genre Namespace Role

primary agent/s

Mandatory administrative metadata required is:

The Identity of the metadata registrant (person or organisation)

URL resolution (current web location of resource)

Metadata record version (version number of the metadata)

Date of metadata declaration

Below is an extended DOI kernel for a journal article as a Work (the abstract creation or publisher's production line item) with identifiers which can link to its printed and digital manifestations (for actual trading purposes): it can be contrasted with the University of Arizona DC metadata template, shown earlier, which is not as precise.

SLIDE 15: DOI JOURNAL WORK + MANIFESTATION METADATA KERNEL

Kernel Extension Value Example

DOI **DOI of Work** DOI e.g. 10.1000/**(W)**131

DOI of 1st digital mfst DOI e.g. 10.1000/**(D)**131

DOI of 1st print mfst DOI e.g. 10.1000/**(P)**131

The Handle System can resolve to any of the locations for any of the manifestations

DOI genre Jnl Article Fixed value for genre

Identifier Work identifier PII or ID If PII allocated by publ

Article identifier SICI 0264- 1615(1999)<13 Journal identifier ISSN,CODEN ISSN 0264-1615

Publication date YYYY 1999

Journal vol.no Number 27

Journal issue no. String 1

Article start page no. String 13

Article start page seq.no Number Mandatory

Title Article title String The DOI system

Journal title String Info Science Studies

Type Physical format e.g., print print

Digital format e.g., HTML pdf

Origination Work origination e.g. **Work: original**

Manifest orig:dig Original **Phys: original**

Manifest orig:phys Derivation **Dig:version suppl incl**

Excerpt

Compilation

Replica

Agent identifier First named author Auth name/s Maree, J A

Additional authors "et al" "et al" flag

Publisher name/s Publ.name/s XYZ University Press

Administrative metadata

Metadata registrant Registr/publ XYZ University Press

URL resolution URL http://www...etc Metadata version Number 1

Date of submission YYYYMMDD 19990610

The heart of the matter here is that all further manifestations or versions of this article/ Work which are registered with the IDF can be linked to the original Work, which then allows them all to be linked to each other (called 'nesting"). This enables a "rich manifestation tree" which can respond to differing business contexts where "reaching an agreement might involve traversing an object's family tree in search of the appropriate object level at which to do business" (Erickson 1999).

The Work identifier is recommended as the declared identifier in a bibliographic citation; DOIs may also link manifestations with their abstracts where publishers continue to supply resource information to their normal A&I services as well.

Academic Press has been the first to implement Level 2 DOI identifiers and metadata for their licensed IDEAL online journal library product (April 1999).

The DOI is not an intellectual property protection scheme by itself; rather, it is seen as a key element in structured online transactions for publishers' products, rather like an item barcode or ISBN. However, the DOI carries the extra functionality of reliable navigable linking which brings publishers closer to their declared objective of integrating content with services.

5: CONCLUSION

The various content-owning and trading industries admit that their various rights administration business modes are not conducive to collectively robust ecommerce, particularly when surveying the fragmented copyright landscape of art, music, cinema, book and journal publishing.

Consensus around a digital architecture of unique object identifiers supported by what Eric Hellman (1999) calls "transactable metadata" forms the beginning of coherent online intellectual property description, protection and trade.

The information industry will be transformed by these initiatives, and libraries, who are significant trading partners but who are also approaching various states of "being digital" in which they too will begin to create information resources or packages for their customers, and possibly for commercial profit, must take note of these developments.

6: ADDENDUM: EMERGING METADATA TOOLS

The starting point for a common understanding between the rights and the descriptive communities was the IFLA analysis of Functional Requirements for Bibliographic Records (FRBR), which was adopted as a common logical model for resource classification. This led to the acceptance of the Resource Description Framework (RDF) as a common syntax model for expressing the semantic content of both INDECS/DOI and DC forms of metadata descriptions; which led naturally to the acceptance of XML as the preferred markup language, since the RDF is best expressed in XML.

These three metadata tools have drawn the DOI/ INDECS and DC communities closer to each other. They are briefly described below:

6.1: THE IFLA FUNCTIONAL REQUIREMENTS FOR BIBLIOGRAPHIC RECORDS

(FRBR) LOGICAL MODEL

The IFLA Study Group on the Functional Requirements for Bibliographic Records produced its final report in 1998. The Group recognised that new forms of electronic publishing and network access to information resources had led to a broad range of user expectations and needs, and that it was necessary to re-examine the relationship between data elements in a resource description and the needs of the user. Of particular interest to the INDECS/DOI and DC communities is the report's high-level classification of resources into 4 possible "states". This is an attempt to separate the abstract concept of a work from its potential many, varied and perhaps variable digital expressions: there is a growing awareness that digital resources cannot be described by rules such as AACR for fixed in-hand copies (Heaney 1998).

SLIDE 16: THE FRBR LOGICAL MODEL

1. WORK: The original creator conceived a **WORK**. (abstract state, e.g.Beethoven 5th Symphony)

DC/ INDECS AGREE: CAN ADD: AGENT, ACTIONS, SUBJECT, OBJECT RELATIONS

2.EXPRESSION: The WORK can be realised in one or more **EXPRESSIONS** (

performance of the Symphony)

DC/ INDECS AGREE:CAN ADD: AGENT, ACTIONS, SUBJECT, OBJECT RELATIONS

3. MANIFESTATION: EXPRESSIONS can be embodied in one or more **MANIFESTATION** (printed score, CD recording, video recording)

DC/ INDECS AGREE: CAN ADD AGENTS, ACTIONS, SUBJECT, OBJECT RELATIONS

4. ITEM: MANIFESTATIONS can be exemplified in many (mass-produced) copies or **ITEMS.** Alternatively only one copy may exist, e.g. a museum piece**.**

It becomes relatively easy to identify a Creation in any of these logical states, and thence to apply appropriate metadata, whether rights-based or bibliographic, to a resource in any of these states, and to cascade the metadata from one state or level to the next; to put it in more recognisable "cataloguing" terms, the FRBR makes it easier to identify logical categories of content types to which metadata might be applied by different communities: the author/ creator can report metadata about the Work (whether it is a novel, musical score or article in a journal), a publisher/copyright owner could report the metadata of a print or digital Manifestation (in whichever Genre), whilst Performance metadata can document the expressions of the work so performed.(It must be remembered that the INDECS project also encompasses communities such as producers of recorded music, cinema, television, theatre).

INDECS and the DC community have agreed that this logical model can be expanded by adding AGENT relationships (persons, corporate bodies) to each of the four Creation categories, as well as ACTIONS (originate, acquire, compile, excerpt); SUBJECT relations (a work is about something, someone) and OBJECT relations (a creation is made of or part of something) can be added to satisfy the descriptive requirements of both the DC bibliographic and the INDECS/DOI rightsholding communities.

6.2: THE RESOURCE DESCRIPTION FRAMEWORK (RDF) SYNTAX

The Resource Description Framework (RDF) was developed by the World Wide Web Consortium (W3C), directed by Tim Berners-Lee, who also pioneered the use of the URL, HTTP and HTML.It is the Consortium's solution for interfacing the various metadata sets which are emerging from a range of information communities (geographic data, government information, scientific statistical research communities, museums) in "silo" fashion, and which are in danger of being unable to communicate with each other.

The RDF does not prescribe a fixed metadata terminology, but provides a consistent framework or metadata architecture which allows different communities (namespace domains) to consistently use their own agreed naming conventions or rules for data, and to structure them so that data elements can be consistently tagged, searched and matched in a standardised way across the differing schemas, especially by search engines or intelligent software agents. The RDF allows communities to focus on defining their data element vocabularies, rather than try to design the structure and syntax for metadata as well.

The basic RDF model requires that statements about a resource are grouped into a 'description element'; the statements classify resources in a triadic model of RESOURCES (things, books, videos) which have PROPERTIES (height, width, country of origin) with VALUES (100mm, published in Spain). This is very similar to the Dublin Core and INDECS models for describing RESOURCES by means of their ATTRIBUTES and ELEMENTS, which is why the Framework is accepted by both these communities. The description element may be part of the web document or may be stored in a separate data resource.

SLIDE 17: DESCRIPTION ELEMENT EXPRESSED IN RDF/ XML FRAMEWORK:

using description elements from DC, DOI and IEEE namespaces/ schemas/ vocabularies

<? xml version = "1.0" ?> **[states this is an XML document]**

<RDF xmlns = "http://w3.org/TR/1999/PR-rdf-syntax **[XML namespace RDF]**

*Search engines can check at the xmlns addresses for the official 'vocabularies' and their element values; and may crossmap to obtain similar values from the different vocabularies: i.e. DOI perhaps defines 'author' differently to DC and to IEEE, but crossmapping will allow search engine to retrieve all as authors.

*XML namespaces mean you don't have to use the old HTML LINK tag to link or refer every element individually to its vocabulary or schema.

xmlns = "http://purl.org/DC#"> [properties in description come from DC n/space]

xmlns = "http://doi.org#"> properties also from DOI namespace]

xmlns = "http://ieee.org#> [properties also from IEEE namespace]

Description begins:

<Description about = "http://dtsc.com.au/report.html"> **[URL identifier of doc]**

<Description about = "http://dx.doi/org/10.1000.92 **[DOI identifier of doc]**

<DC:title> = Metadata management</DC:title> **[property defined in DC schema]**

<DC:creator> = Jean Maree </DC:creator> **[property defined in DC schema]**

<DC:date> = 1999-05-21 </DC:date> **[property defined in DC schema]**

<IEEE:copyright = IEEE 1999 </IEEE copyright [defined in IEEE schema]

<DC:subject> = metadata, intellectual property </DC:subject> **[defined in DC schema]**

</Description>

Description ends

The governing authority for the properties of each schema or vocabulary is shown in the XML defined (discussed below) namespace mechanism in a description, which points to that official source for the definitions of the properties that were used to create the metadata. The RDF infrastructure can also support the combination of attributes from different registries (DOI, DC, IEEE), as shown in slide 16.

By consulting the namespace or schema authorities directly for definitions, and using resolution rules for conversion or equivalences, search engines will be able to map, translate or correlate these data tags, decide which are similar or dissimilar, extract the content value from those tags which are similar, and produce meaningfully merged search results from RDF structured descriptions.

The DC use of RDF is shown here; the INDECS use of RDF is still under development.

SLIDE 18: DUBLIN CORE and DC QUALIFIERS COMBINED WITH IMAGINARY VCARD ELEMENTS AND EXPRESSED IN RDF/XML

?**xml version='1.0**'?>

<**rdf:RDF xmlns:rdf**="http://www.w3.org/1999/02/22-rdf-syntax-ns#"

LISTING THE RELEVANT NAMESPACES

xmlns:dc = "http://purl.org/**dc/elements**/1.0/"

xmlns:dcq = "http://purl.org/ **dc/qualifiers**/1.0/"

xmlns:vcard = "http://www.imc.org/ **vcard**/3.0/">

DESCRIBING THE AUTHOR/ ILLUSTRATOR WITH DC + DCQ ELEMENTS

<**rdf:Description rdf:about** = "http://**doc**">

<**dc:creator**>

<**rdf:Description**>

<**rdf:type rdf:resource** = "http://purl.org/dc/terms/1.0/**creator/class/Person**"/>

<**dcq:creatorType rdf:resource** =
"http://purl.org/dc/terms/1.0/**creator/type/Illustr**"/>

<**rdf:value rdf:resource** = "http://411.com/**JoeSmith**"/>

</**rdf:Description**>

</**dc:creator**>

</**rdf:Description**>

QUALIFYING THE AUTHOR / ILLUSTRATOR WITH "VCARD" ELEMENTS

<rdf:Description rdf:about = "http://411.com/**JoeSmith**">

<**vcard:fn**> Joe Smith </vcard:fn>

<**vcard:email**> joe@my.com </vcard:email>

<**vcard:org**> My Company, Inc.</vcard:org>

</**rdf:Description**>

</**rdf:RDF**>

6.3: eXTENSIBLE MARKUP LANGUAGE (XML)

HTML is a presentation orientated or document display markup language, mainly designed to arrange a document for visual display in a browser, showing paragraphs, bold type, colour, font size, hyperlinks, etc; it uses a fixed repertoire of specifically named tags and tag

meanings, such as P for Paragraph, H1 for Heading1. For instance, a Web browser will interpret any text between the <H1> and the </H1> tags as a major heading, and will display it in large bold type. This does not tell us anything about the content of the document. The only tag which permits inclusion of description and retrieval terms to indicate the information content of a document is the <META> tag.

SLIDE 19: HTML DOCUMENT TAGGING

<P>

This is the first paragraph

</P>

<P>

This is the second paragraph

These are some special effects:

This is a word in <I>italics</I>.

This is a word in <B>bold</B>.

Here is an in-lined GIF image: <IMG SRC = "myimage.gif">.

</P>

<P>

This is the third paragraph, which demonstrates links. Here is

a hypertext link from the word <A HREF = "subdir/myfile.html">foo</A>

to a document called "subdir/myfile.html".

</P>

<H2>A second-level header</H2>

Here is a section of text that should display as fixed-width font

etc

A new and more flexible markup language is XML, which can accommodate documents and metadata more easily than HTML by separating content and presentation and by relegating the presentation aspect of the document to the use of various stylesheets using XSL (eXtensible Stylesheet Language). XML user communities may then create their own descriptive tags for document content. " XML is intended to take Web pages beyond the basic display capabilities of HTML and transform documents into contextual objects that can perform more useful electronic feats...if XML becomes the universal format for structuring data,..everything from classified ads to news articles to recipes to search engines will be more orderly, and applications like browsers can manipulate and display these documents in more useful ways " (Oakes 1999). Note that XML is used for expressing the entire document, and not just for its metadata label or description.

SLIDE 19: XML DOCUMENT TAGGING

1. TO SCHEDULE INTERACTIVE MEDIA PRESENTATION ON WEB

<par>

<audio src = "narrator.au"/>

</par>

(this example plays a soundtrack, and after 2 seconds displays an aerial photograph for a further 3 seconds)

2. DISPLAYING MUSIC NOTATION

<?xml version = "1.0">

<DOCTYPE sheetmusic SYSTEM "music.dtd">

<sheetmusic>

<musicrow size = "two>

<entrysegment>

<entrypart cleff = "treble" rhythm = "threequarter" position = "one">

<molkruis level = "zero" name = "f"

notetype = "sharp"/>

</entrypart></entrysegment>

The main web browsers, Microsoft and Netscape, are beginning to support and incorporate XML, and several XML authoring tools have been developed, e.g. XML<PRO> from Vervet Logic and the ADEPT Editor 7.0 from ArborText. It is certain that "XML will advance the Web as a medium for document delivery, publishing, workflow management and data exchange" (Funke: 63). Particularly exciting XML developments are tagging schemes for music notation and mathematical symbol markups.

Implementations of the RDF syntax can be expressed by XML for the exchange of metadata across different namespace domains, as long as a namespace identifier is prefixed to each metadata element, e.g. <doi.title> or <dc.author>, thereby allowing the construction of equivalences between schemas.

****

NOTE 1: WEB SEARCH ENGINES AND METADATA

It should be noted that the universal web search engines do not currently support or "harvest" DC metadata, as DC can only be fully searched if it is in the syntax recommended for a particular engine. Alta Vista and HotBot tend mainly to index data in the KEYWORD and DESCRIPTION qualifiers in the META tag of an HTML document (Powell 1997), where an author

may not have used DC-structured metadata, but has simply stated the following in freetext :

<META NAME= "description" CONTENT = "A paper on metadata"

<META NAME= "keywords" CONTENT = "DOI, Dublin Core, INDECS"

It is hoped that developers of these search engines will eventually modify their software to enable searching in DC-described <META> tag fields as well, so that the search engines can retrieve, parse, match and filter these embedded tags to deliver more focused search results.

Search engine developers have not been as eager to participate in metadata developments as the browser designers (Sullivan 1997), reasoning that it is not yet used widely enough to merit modification of their machines from free-text to structured searching. Diann Rusch-Feja (1998) also mentions that the universal search engines pride themselves on the number of hits which they can produce, and therefore might logically not be interested in reducing numbers of hits to obtain search precision, especially when this would limit their advertising facility . Another problem is that unscrupulous web page designers could deliberately manipulate the ranking and indexing of their pages by means of keywords.

Some subject-oriented servers of individual information communities are DC-compatible, e.g. the Australian Metadata Search Engine and the Nordic Web Index.

****
*

NOTE 2: METADATA IS NOT THE SAME AS CATALOGUING

It should perhaps immediately be noted that metadata is not merely the equivalent of library cataloguing. In a paper delivered at the 1998 IFLA Conference, Stefan Gradmann of PICA reproached those who promote the idea that "a library catalogue record could be described as metadata", because it too is ‘data about data’, therefore ‘carry on cataloguing’ as if nothing has happened (Gradmann 1998).

Gradmann notes that:

1. The 'user’ of metadata is different. Metadata is chiefly meant to be machine understandable, so that software agents can understand and process it in automated searches and merge it in results displays.

2. The purpose of metadata is different. Metadata is meant to enable dynamic resource discovery more than it is meant to facilitate classic resource description.

3. The production source of metadata is different. Metadata is not necessarily produced by professional cataloguers, but should ideally be produced at the source, i.e. by the author, a publisher or a third party such as an abstracting service.

4. Context and end-user behaviour dictate the appropriate metadata, such as specific kinds of resources, specific kinds of user groups or specific kinds of usage. Library cataloguing is more generic and is less context sensitive.

5. Technical context is also important, as the metadata record is usually embedded into the resource describes, and it usually contains and points to the location of the online resource. A library catalogue record attaches an item/ shelf record to a free-standing bibliographic record; this sometimes requires staff assistance to ‘point’ the user at the book . Even where the metadata is a distinct data object itself, possibly stored in a repository, it is always technically linked to its object to a high degree, and metadata producers must therefore take continual cognisance of this and of changing technological standards to maintain links.

Gradmann concludes that metadata belongs to a different production paradigm, and that cataloguers should prepare to redefine their roles within these changing contexts, although he notes that their expertise with authority forms, controlled vocabularies and consistency or quality assurance for 'trusted resource descriptions', which may yet prove to be their leveraging strengths in the developing metadata industry.

****

7: **BIBLIOGRAPHY**

Alrashid, T et al 1998 [online]: Safeguarding copyrighted contents: digital libraries and intellectual property management. *D-Lib Magazine, April*.[cited 12/05/99]

Available from World Wide Web: <http://www.dlib/april98/04barker.htm>

Bearman, D et al 1999 [online]: A Common model to support interoperable metadata; progress report on reconciling metadata requirements from the Dublin Core and INDECS/DOI communities. *D-Lib Magazine, 5 (1)*. [cited 19/01/99]

Available from World Wide Web: <http://dlib.org/dlib/january99/bearman/01bearman.html>

BIC/ EDITEUR/NISO Rights Metadata Working Party 1998 [online]: Interim report 6/2/98 [cited 15/05/99]

Available online: <http://www.bic.org.uk/rights/html>

Brodie, M 1997 [online]: In the company of strangers: challenges and opportunities in metadata implementation; paper delivered at the Conference on Beyond the Beginning: the Global Digital Library, 1997 . Washington: Coalition for Networked Information. [cited 11/05/99]

Available from World Wide Web: <http://www.cni.org/regconfs/1997/ukoln-content/repor~29.htm>

Cathro, W 1997 [online]: Metadata: an overview. Paper delivered at the Australia Standards Seminar "Matching discovery and recovery", August 1997. [cited 08/08/98]

Available from World Wide Web: <http://www.nla.gov.au/nla/staffpaper/cathro3.html>

Causton, L 1998 [online]: Identifying and describing web resources. *Interactive Electronic Publishing*. European Commission DGXIII/E-4. [cited 25/03/99]

Available from World Wide Web: <http://elpub.org/html/webres.htm>

Davis, J R & Klavans, J L 1996[online]: Workshop report: the technology of terms and conditions. *D-Lib Magazine, June* .[cited 12/05/96]

Available from World Wide Web: <http://dlib.org/june97/06davis.htm>

De Niet, M 1998 [online]: DONOR WP2 Metadata generator; functionality of metadata generators. Version 1.0.[cited 17/03/99]

Available from World Wide Web: <http://www.konbib.nl/coop/donor/rapporten/generatorspec.htm>

Dempsey, L 1996 [online]: Meta detectors. *Ariadne, issue 3.* [cited 04/05/99]

Available from World Wide Web: <http://www.ariadne.ac.uk/issue3/metadata>

Dempsey, L & Weibel, S 1996 [online]: The Warwick Metadata Workshop: a framework for the deployment of resource description. *D-Lib Magazine, July- August.*[cited 19/03/99]

Available from World Wide Web:

<http://mirrored.ukoln.ac.uk/lis-journals/dlib/dlib/dlib/july96/07weibel.htm>

Dublin Core Metadata Initiative 1999 [online]: Guidance on expressing the Dublin Core within the Resource Description Framework (RDF). Draft proposal edited by E. Miller et al. [cited 27/05/99]

Available from World Wide Web:
<http://www.ukoln.ac.uk/metadata/resources/dc/datamodel/WD-dc-rdf >

Erickson, J S [online]: The role of metadata supply chains in DOI-based, value-added services. [cited 20/05/99]

Available from World Wide Web: <http://www.ybp.com/yps/papers/doiservicesapr99.htm>

Funke, S 1998: XML content finally arrives on the Web. *Searcher: the magazine for database professionals, 6, (9)*: 63- 68.

Gladney, H M 1998 [online]: Safeguarding digital library contents and users; a note on universal unique identifiers. *D-Lib Magazine, April*. [cited 12/05/99]

Available from World Wide Web: <http://www.dlib/april98/04gladney.htm

Gradmann, S 1998 [online]: Cataloguing vs. Metadata: old wine in new bottles? paper presented at the 64th IFLA General Conference, August 16- 21, 1998. [cited 22/09/98]

Available from World Wide Web: <http://www.ifla.org/IV/ifla64/007-126e.htm>

Heaney, M 1998: Models, materials and moments; the model of AACR. *Catalogue & Index; periodical of the Library Association Cataloguing and Indexing Group, 129*: 1-3

Hellman, E 1999 [online]: [Discuss-DOI mailing list]. Email dd 26 May 1999.

INDECS 1999 [online]: Website [cited 31/05/99]

Available from World Wide Web: <http://www.indecs.org>

International Federation of Library Associations. IFLA Study Group on the Functional Requirements for Bibliographic Records 1998 [online]: Functional requirements for bibliographic records. Final report. Munchen: Saur.[cited 04/05/99]

Available from World Wide Web: <http://www.ifla.org/VII/s13/projects.htm>

International DOI Foundation 1999 [online]: About the DOI. Washington, DC: IDF. [cited 12/05/99]

Available from World Wide Web: <http://www.doi.org/about_the_doi.htm>

International DOI Foundation 1997 [online]: Guide to using Digital Object Identifiers, for creators, publishers and information providers. Washington, DC: IDF. [cited 08/06/98]

Available from World Wide Web: <http://www.doi.org/guidebook/guidebook.html>

Kahn R & Wilensky R 1995 [online]: A framework for distributed digital object services. Reston: Corporation for National Research Initiatives. [cited 23/04/99]

Available from World Wide Web: <http://www.cnri.reston.va.us/cstr/arch/k-w.htm>

Lynch, C 1998 [online]: The Dublin Core descriptive metadata program: strategic implications for libraries and networked information access. *ARL Newsletter 196.* Washington: Association of Research Libraries. [cited 11/-05/99]

Available from World Wide Web: <http://www.arl.org/newsltr/196/dublin.htm

Lynch, C 1997 [online]: Searching the Internet. *Scientific American, March*. [cited 11/05/99]

Available from World Wide Web: <http://www.sciam.com/0397issue/0397lynch.htm

Lyons, P 1997 [online]: Managing access to digital information: some basic terminology issues. *ASIS Bulletin, Jan/Dec*. [cited 20/05/99]

Available from World Wide Web: <http://www.asis.org/Bulletin/Dec-97/lyons.htm>

Maree, J 1998 [online]: The Digital Object Identifier system and its applications in the online information industry; an overview for Southern African information workers. *Musaion, XVI (2)*: 69- 81.

Available from World Wide Web: <http://www.unisa.ac.za/library2/afdeling/tdienste/doi.html>

Miller, E 1998 [online]: An introduction to the Resource Description Framework. *D-Lib Magazine, May 1998*. [cited 23/04/99]

Available from World Wide Web: <http://www.dlib.org/dlib/may98/miller/05miller.htm>

Miller, P 1997 [online]: Metadata for the masses.*Ariadne (5)*. [cited 03/05/99]

Available from World Wide Web: <http://www.ariadne.ac.uk/issue5/metadata-masses>

Newton, J J 1998 [online]: Using levels of abstraction to name data elements. *Data Administration Newsletter (TDAN) 7.0*. [cited 03/05/99] Washington: National Institute of Standards and Technology (NIST).

Available from World Wide Web: <http://tdan.com/i007ht01/html>

Oakes, C 1999 [online]: Building a self-actualised Web. *Wired News, 9 March*. Wired Digital Inc. [cited 11/03/99]

Available from World Wide Web:
<http://www.wired.com/news/print_version/technology/story/18356.html>

Oppenheim, C 1998: The legal and regulatory environment for electronic information. 3rd ed. Tetbury: Infonortics Ltd.

Paskin, N 1998a [online]: The Digital Object Identifier initiative: current position and view forward. DOI discussion paper. Version 3, August 1998. Washington: International DOI Foundation. [cited 10/09/98]

Available from World Wide Web: <http://www.doi.org/resources.htm>

Available from World Wide Web: <http://dx.doi/org/10.1000/92>

Paskin, N 1998b [online]: The Digital Object Identifier initiative: metadata implications. Version 2, November 1998. Washington: International DOI Foundation. [cited 10/01/99]

Available from World Wide Web: <http://www.doi.org/resources.htm>

Paskin, N 1999a [online]: DOIs used for reference linking. Version 1.0. Washington: International DOI Foundation. [cited 30/03/99]

Available from World Wide Web: <http://www.doi.org/resources.htm>

Paskin, N 1999b [online]: DOI: current status and outlook, May 1999. *DLib Magazine*, May. [cited 20/05/99]

Available from World wide Web: <http://www.dlib.org/dlib/may99/05paskin.htm>

Paskin, N & Rust, G 1999 [online]: The Digital Object Identifier initiative: metadata implications. Version 3: February 1999. Washington: International DOI Foundation. [cited 20/02/99]

Available from World Wide Web: <http://www.doi.org/resources.htm>

Partyka, J 1999 [online]: W3C's Berners-Lee urges agent-readable Web sites. *InfoWorld Electric*. [cited 29/03/99]

Available from World Wide Web:

<http://infoworld.com/cgi-bin/displayStory.pl?990325.eibernerslee.htm>

Powell, A 1997 [online]: Dublin Core management. *Ariadne, issue 10*. [cited 11/05/99]

Available from World Wide Web: <http://www.ariadne.ac.uk/issue10/dublin>

Rusch-Feja, D [online]: Metadata: standards for retrieving WWW documents (and other digitised and non-digitised resources). *Library and Information Services in Astronomy III; ASP Conference series, (153)*. San Francisco: Astronomical Society of the Pacific. [cited 11/05/99]

Available from World Wide Web: <http://www.stsci.edu/meetings/lisa3/ruschfejad.htm>

Rust, G 1998a [online]: Metadata: the right approach. *D-Lib Magazine, July-August*. [cited 30/07/98]

Available from World Wide Web: <http://www.dlib.org/july98/rust/07rust.html>

Rust G 1998b [online]: The fire and the rose; an integrated model for descriptive and rights metadata. London: Data Definitions. [cited 20/09/98]

Available from World Wide Web: <http://www.bic.org.uk/rights.htm>

St. Pierre, M & LaPlant, W P 1998 [online]: Issues in crosswalking content metadata standard [cited 05/03/99]. Washington: National Information Standards Organization.

Available from World Wide Web: <http://www.niso.org/crsswalk.htm>

Sullivan, D 1997 [online]: The new meta tags are coming- or are they? *The Search Engine Report, Dec 4*. [cited 11/05/99]

Available from World Wide Web: <

http://www.searchenginewatch.internet.com/sereport/9712-metatags.htm>

Van der Werf-Davelaar, T 1999 [online]: Identification, location and versioning of web-resources; URI discussion paper. Version 1.0. [cited 17/03/99]

Available from World Wide Web: <http://www.konbib.nl/donor/rapporten/URI.htm>

World Wide Web Consortium 1999 [online]: Resource Description Framework (RDF) schema specification; W3C proposed recommendation 3 March 1999 [cited 31/03/99].

Available from World Wide Web: <http://www.w3.org/TR/PR-rdf-schema>

Weibel, S 1999 [online]: The state of the Dublin Core Metadata Initiative April 1999. *D-Lib Magazine, 5 (4).* [cited 19/04/99]

Available from World Wide Web: http://www.dlib.org/dlib/april99/04weibel.htm

---

Copyright © 2001