An Analytical Model of Collections and their Catalogues

A study carried out by

Michael Heaney
University Library Services Directorate
University of Oxford

on behalf of the
UK Office for Library and Information Networking
with support from
OCLC

Third Issue, Revised
Oxford, 14 January 2000

Contents

0. Introduction

0.1 This work arises from the UKOLN proposal to the Research Support Libraries Programme *Collection Development: Study, Recommendation, Specification*, which undertakes to 'refine our current approach based on a more thorough modelling of collections and their catalogues', and to validate the approach by working with RSLP projects and others to describe their collections. The elaboration of the model itself has been carried out with financial support from OCLC.

0.2 The model has been designed without regard to any specific implementation. UKOLN proposes to translate the model into a schema, and from there to construct a demonstrator implementation, in subsequent phases of the project. The need to reflect the complexity which underlies collection description has led to a multidimensional model, and some of the possible vehicles for implementation may not support such a structure fully. Schemes such as RDF and XML may provide a richer implementation, with secondary mappings to simpler standards such as HTML; but it is inevitable that some aspects of the structure will be lost in such mappings. It is hoped that the model is comprehensive enough to clarify the differences between those aspects of any implementation which truly reflect the reality of collection description and those which merely derive from the structure of the implementation mechanism.

0.3 Collection description is such a broad descriptive term that it is worth saying something about the intended scope of the model. Although it has its origin in the RSLP programme, many of whose results will be digital resources of one kind or another, the model is not restricted to the description of digital collections. It is intended that the model should be applicable to physical and digital collections of all kinds, including library, art and museum materials, and is by no means applicable only to the resources of large research libraries. Collection description itself may take a variety of forms, and the model makes no presumption about the format of such a description.

0.4 The model is aimed in the first instance at those responsible for the development of collection descriptions. It is also a general contribution to the debate about metadata in the digital age. As described above, its initial use will be to inform the construction of a demonstrator to which all relevant RSLP projects can feed information. With the model as its base, the demonstrator is intended to be appropriate for and hospitable to their requirements for collection description. In terms to be developed below (see section 5.5 and 6), the demonstrator will accommodate Unitary Finding-Aids for the collections.

0.5 Although the primary purpose of this model is to illumine the process of resource discovery by users, collection description also serves collection management purposes, particularly in discharging an institution's curatorial responsibilities.

1. The information landscape

1.1 The information landscape can be seen as a contour map in which there are mountains, hillocks, valleys, plains and plateaux. A large general collection of information – say a research library – can be seen as a plateau, raised above the surrounding plain. A specialized collection of particular importance is like a sharp peak. Upon a plateau there might be undulations representing strengths and weaknesses.

1.2 The scholar surveying this landscape is looking for the high points. A high point represents an area where the potential for gleaning desired information by visiting that spot (physically or by remote means) is greater than that of other areas. To continue the analogy, the scholar is concerned at the initial survey to identify areas rather than specific features – to identify rainforest rather than to retrieve an analysis of the canopy fauna of the Amazon basin. This model attempts to characterise that initial part of the process of information retrieval.

1.3 The landscape is, however, multidimensional. Where one scholar may see a peak another may see a trough. The task is to devise mapping conventions which enable scholars to read the map of the landscape fruitfully, at the appropriate level of generality or specificity.

1.4 The IFLA study *Functional Requirements of Bibliographic Records* identifies (pp.8-9) four functions of records, progression through which may be seen as constituting a successful traverse of

the information landscape and the attainment of one's goal. These are:

To find   i.e. to provide access points by which information can be found

To identify   i.e. to describe something so as to enable users correctly to interpret records retrieved

To select   i.e. to provide a means for users to choose from among the identified records

To obtain   i.e. to acquire the identified materials.

1.5 The first two of these activities are associated with the traditional areas of catalogue codes, access and description. The relations they embody are characteristically static or at least persistent. A static model may adequately represent them – they are the map of the landscape. The second two reflect the more active operations involved in retrieving and using information; they are transactional in nature, and a dynamic or event-driven model may be more appropriate for them – they represent attempts to use the map to reach the areas of interest.

1.6 The model will attempt to encompass the first two activities. There are, however, many links to be made between all the elements in the process of obtaining information, and these links may be expressed reciprocally. Determining and describing the part of a link which may be embodied in the model inevitably determines the nature of the complementary half, though the objects at the other end may not be described fully, or at all.

2. Collections

2.1 The preliminary work done by UKOLN with respect to collection description identified 'collection' as encompassing the following types of entity:

> Internet catalogues (e.g. Yahoo)
> Subject gateways (e.g. SOSIG, OMNI, ADAM, EEVL, etc.)
> Library, museum and archival catalogues
> Web indexes (e.g. Alta Vista)
> Collections of text, images, sounds, datasets, software, other material or combinations of these (this includes databases, CD-ROMs and collections of Web resources)
> Collections of events (e.g. the Follett Lecture Series)
> Library and museum collections
> Archives
> Other collections of physical items
> Digital archives

2.2 In fact these types can be categorised into those that are collections of entities (e.g. books) or of derived representations of entities (e.g. photographs of pieces of sculpture) on the one hand, and those that are collections of information about such entities. (The type 'Collections of events' is problematical, unless what are intended are collections of records of events.) This study refers to a collection of entities as a 'Collection' and to a collection of information about such entities as a 'Collection-Description'.

2.3 Some types of Collection-Description can themselves be seen as Collections, in this case of metadata rather than primary

information. The Creators, Producers &c of the secondary Collection will not necessarily be those of the Collection it catalogues, however. Moreover, the secondary Collection can have its own recursive Collection-Description: an Indexing Finding-Aid may have a Unitary Finding-Aid.

2.4 This study also uses 'Collection-Description' to encompass both intellectually created resources and passive assemblages of data such as those gathered by robotic search engines.

2.5 The primary object of the current exercise is to enable the creations of Unitary Finding-aids for Collections and Collection-Descriptions of all types.

2.6 The model says nothing explicit about the size of a Collection. It is possible to envisage a 'Collection' consisting of one Item. Where an institution can choose between different degrees of aggregation in determining what are its Collections, there is no structure inherent in the model which requires or predisposes a particular level of aggregation. The institution should base its choices on its own pragmatic grounds, such as the level of detail required to make explicit those elements of the Collection-Description which the institution deems to be useful or necessary for the purposes of resources discovery or collection management (i.e. should adopt a 'functional granularity' approach).

3. Relevant work

3.1 As a preliminary to a detailed elaboration of the Entities and Relationships identified, comparisons were made between the Entities in the present model and the concepts identified in other current work in the field. There has not been sufficient time to examine all the potentially relevant schemes, and it was decided to sample a variety of approaches in order to identify the requirements of different interest groups. Those used are:

> *Dublin Core;*
> The preliminary analysis of Collection Description by UKOLN;
> *ROADS Cataloguing Guidelines* (Note: the Cataloguing Rules were used as a guide rather than the experimental Collection Template, as the latter merely embodies the preliminary analysis cited above);
> *MODELS Profile Interoperability Sub-Set* (the subset of the MODELS profile originated by M25 Link and developed by the Interoperability Focus Group);
> *Functional Requirements of Bibliographic Records;*
> *General International Standard Archival Description;*
> *Towards a European Standard for Manuscript Description:the MASTER project.*

The full listing of concepts identified is given in Appendix A. for full references see Appendix B.

3.2 The *Analytic Model of AACR* was not used as that analyses catalogue rules rather than real-world entities. A complete listing of

the entities identified in *Functional Requirements of Bibliographic Record* would be too lengthy for a summary table, so only the salient elements are identified below. A complete technical specification is not available at the time of writing for the MASTER project; the elements have been identified from the summary description and examples available in the published literature. In some cases a single element in one of the source schemes has been mapped to more than one entity in the current scheme.

4. The model
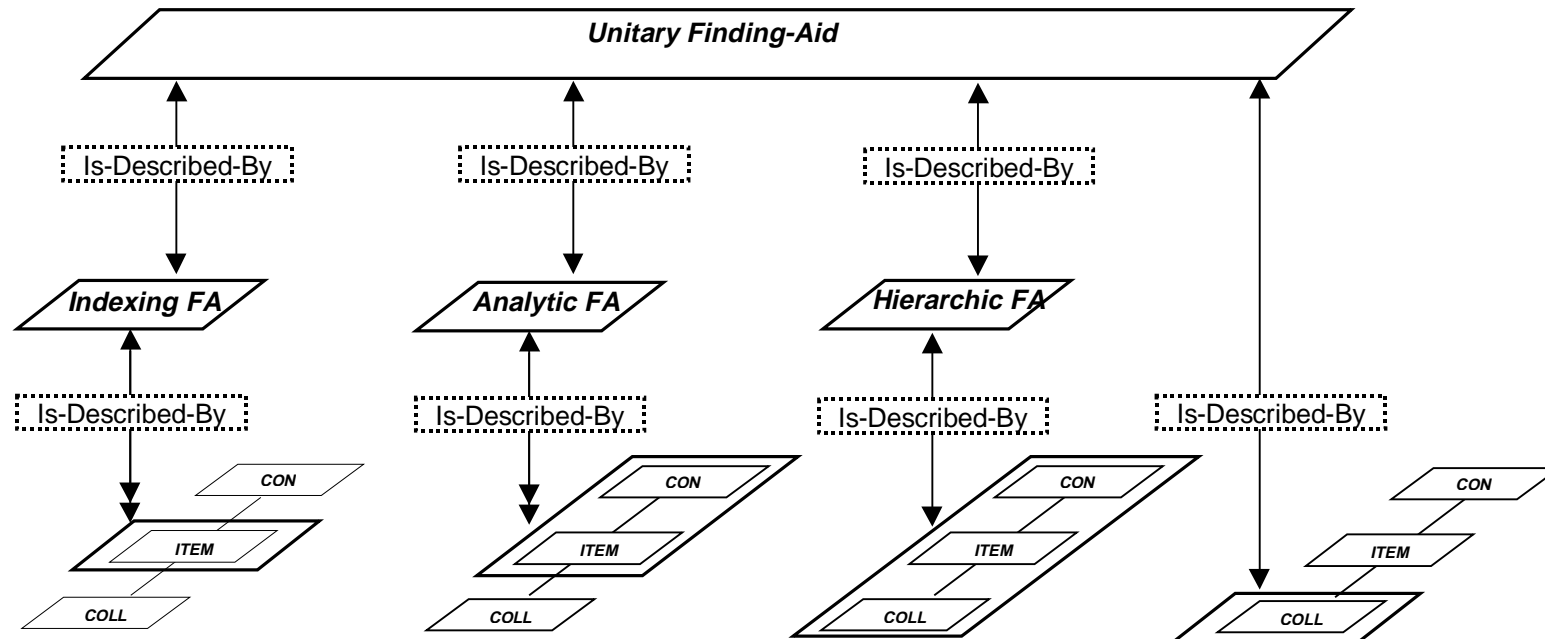
4.1 Collections
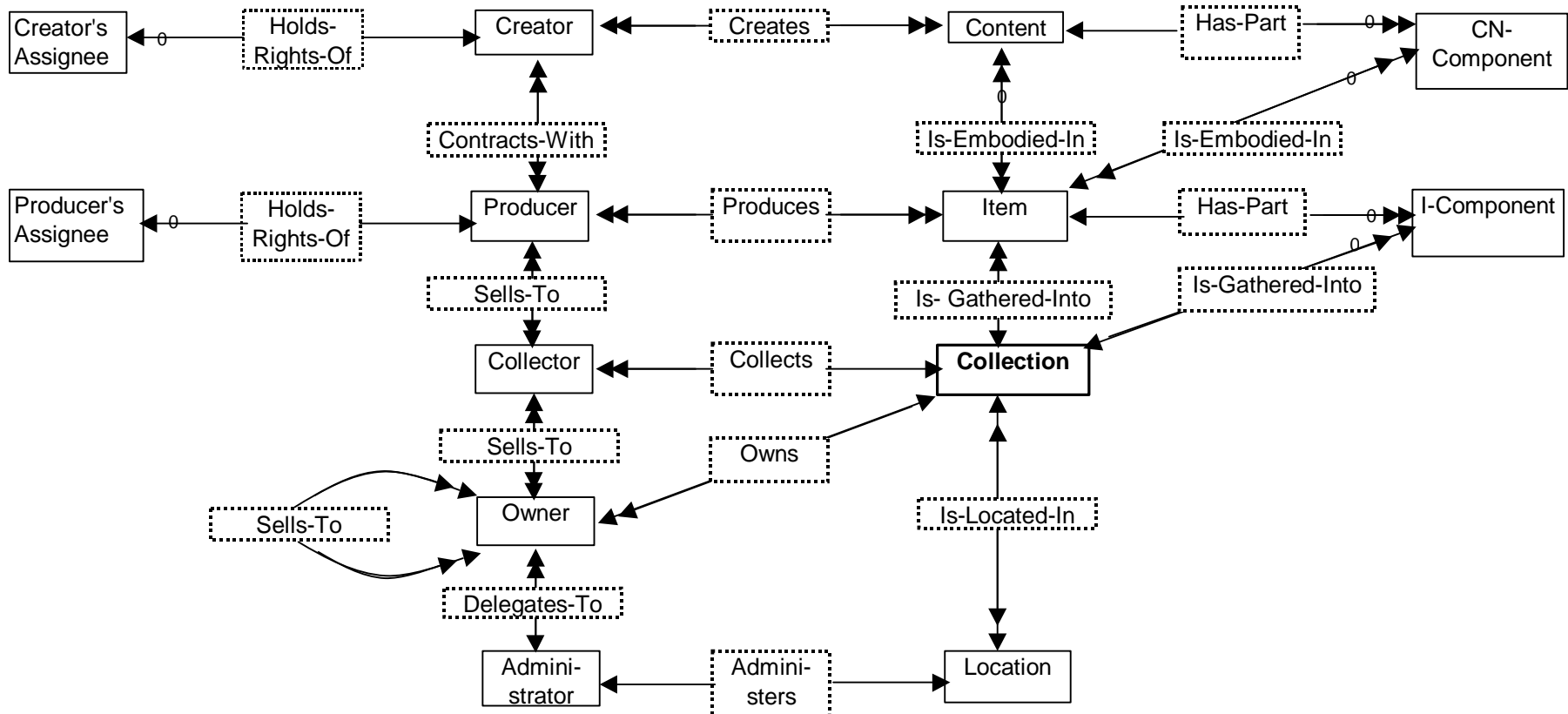
Boxes with solid lines indicate Entities. Boxes with broken lines, and their associated arrowed lines, indicate Relationships. A double arrowhead indicates that multiple instances of the relevant entity may occur (e.g. one Owner may own many Collections but each Collection may have only one Owner). Optional Relationships are indicated by a zero (0) on the relevant arrow.

4.2 Relation between Collections and Collection-Description

The diagram shows four instances of three entities (Content, Item and Collection) from the Collection model. Each instance is linked by the Is-Described-By Relationship to an instance of a Collection-Description. Three of the Collection-Descriptions are also linked to the fourth. (See discussion in Section 6.)

4.3 Collection-Descriptions
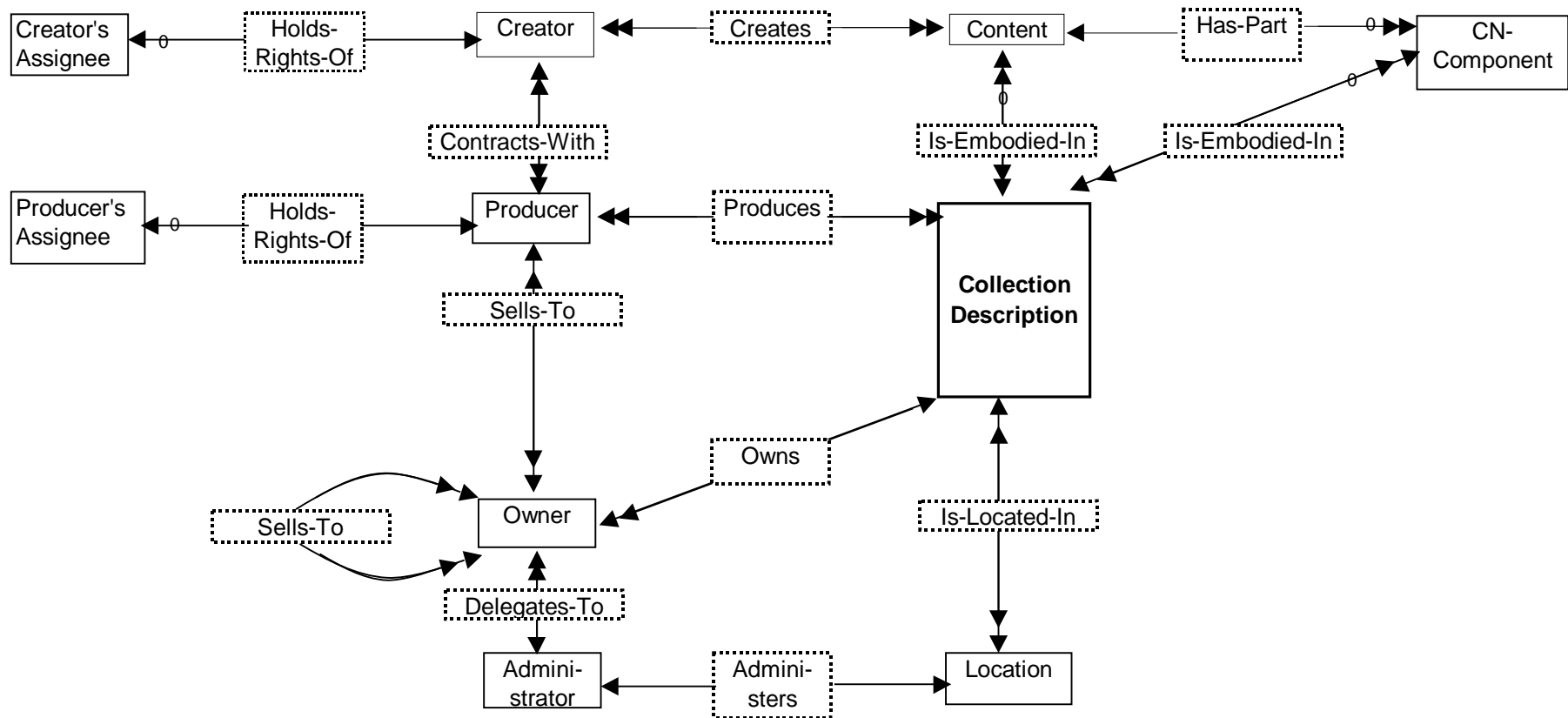
The model for Collection-Descriptions is analogous to that for Collections; the difference lies in the fact that a Collection-Decsription is equivalent to a single Item, so the Entities Collector and Collection, and the Relationship Collects, are not needed.

4.4 The model depicted in sections 4.1-4.3 is a single-record model. Any one realisation of the model will analyse a single collection and a single finding aid. This is because many of the entities may be different in the different instances of the model. For example, the publishing, copyright and ownership status of the microform edition of the Thomason Tracts are different from those of the original collection.

5 Definitions and Attributes of the Entities and their Relationships

The following classes of Entity and Relationship are identified in the current model. In the time available for the composition of this report it has not been possible to develop the description to the next level down, i.e. the Data Elements which comprise each Attribute. For the more complex Attributes the description is not exhaustive, and further work is needed to make it so.

5.1 Entities: Objects

The Class of Entities 'Objects' consists of naturally occurring or created entities incapable of action.

5.1.1 Content

An intellectual creation, without reference to any instantiation of it.
> *Attribute:* Title
> *Attribute:* Description

Note: flagged by ISAD as Scope & content/ Abstract.
Attribute: Subject
>> *Subattribute:* Concept/Keywords
>> *Subattribute:* Object
>> *Subattribute:* Name
>> *Subattribute:* Place coverage
>> *Subattribute:* Date coverage

Note: Where a particular subject or classification scheme is used, this may be incorporated as a data element of the Subattribute (e.g. Subattribute: Concept – Data element: LCSH term); but if the structured scheme is available in a form which can be integrated with the present model, then subject may be modelled as a separate entity within the model).
> *Attribute:* Date
> *Attribute:* Identifier

Note: Flagged by MODELS as Standard Number.
Subtype: Text

Note: Subtype is flagged as "Category" by ROADS.
>> *Attribute:* Language

Subtype: Cartographic
>> *Attribute:* Locus (co-ordinates, equinox &c)

Subtype: Music notation
>> *Attribute:* Notation system

Subtype: Sound material
>> *Attribute:* Type of sound [music/speech/. . .]
>> *Attribute:* Language
>> *Attribute:* Medium of performance

Subtype: Graphic
>> *Attribute:* Medium

Subtype: Audio-visual
    Attribute: Medium of performance
Subtype: Realia
    Attribute: Medium

5.1.2 Item

The concrete (incorporating physical and electronic) realisation of Content.
Note: In so far as this analysis is concerned with collections, the entities Content and Item will be considered only to the extent that their types and attributes impinge upon Collection Description. In the vast majority of cases, too, the Items will coincide with what *FRBR* calls Items, not Manifestations. 'Item' has been chosen as the most neutral term in preference to other terms which have been used such as 'Document' or 'Document-like Object'. 'Item' can most easily embrace all of the concepts of physical and electronic, text and non-text, and human and natural creations.
    Attribute: Title
    Attribute: Description
    Attribute: Format (paper or paper substitute; 3D data object; artwork; photographic; disc; tape; computer chip; material)
    Attribute: Date
    Attribute: Version (edition, issue &c.)
    Attribute: Identifier
    Note: Identifier here is an inherent Item Identifier such as ISBN, not a locator such as call number which relates an Item to the Collection (see Relationship Is-Gathered-Into)

    Attribute: Physical Characteristics (extent; dimensions; containing object)

5.1.3. Collection

An aggregation of physical and/or electronic Items.
    Attribute: Title
    Attribute: Description
    Attribute: System of Arrangement
    Attribute: Subject
    Note: For Subattributes see under Content.
    Attribute: Strength
    Attribute: Date
    Attribute: Identifier
    Attribute: Physical Characteristics
Note (1): Some attributes of a Collection arise from the aggregation of the attributes of its constituent Contents and Documents. For example, the Creators of a Collection are the sum of the Creators of each piece of Content; the Date of a Collection is (a) the range of the Dates of the individual Contents and (b) the range of the individual Documents. In addition, however, there is a Date representing the date of accumulation of the Collection as a collection. Similarly, the subject of a Collection need not be the same as the subject of the Contents (e.g. the subject of a Collection of bindings is the binding of the items, not the subject of the Content of the items); a Collection may have Physical Characteristics additional to those of the documents (e.g. prints kept in guardbooks). The Attributes given above are additional to those inherited from Content and Document, although they may share the same names.

Note (2): Some Attributes given above (Description, Strength) are not inherent to Collections but are commonly assigned to them by institutions preparing metadata and are incorporated here to accommodate these practices.
Note (3) Where a particular scheme is used to characterise an Attribute , this may be incorporated as a data element of the Attribute (e.g. Attribute: Strength – Data element: Conspectus term, or Data element: National Register of Collection Strengths Assessment).

5.1.4 Location

The place (identified physically or electronically) where a Collection is held.
Note: It is important to distinguish between the place and the institution responsible for the place; the latter is represented in this model by the term Administrator.
> *Subtype:* Physical repository
>> *Attribute:* Place [Country, city, building]
>> *Attribute:* Identifier
> *Subtype:* Electronic repository
>> *Attribute:* Site
>> *Attribute:* URL

5.1.5 CN-Component

Part of an intellectual creation as defined by Content.
Note: A component part of Content has the same Attributes as Content

5.1.6 I-Component

Part of a concrete realisation of Content, as defined by Item.
Note: A component part of a Item has the same Attributes as an Item

5.2 Entities: Agents

The Class of Entities 'Agents' consists of personal or corporate entities capable of action, and whose relationships with Objects or with other Agents may involve rights of one kind or another.

5.2.1 Agents can be persons or corporate bodies (in the AACR sense, including conferences, events &c). *FRBR* Figure 3.2 (p.14) lucidly depicts that the same class of entities is responsible for successive aspects of the processes and products which it is the aim of this model to characterise. For the purposes of this model these are separated into their distinctive roles. However, in any particular instance of a Collection the same person or body may fill the role of more than one of these roles.

5.2.2 Creator

An Agent responsible in some way for the existence of the intellectual Content of an Item.
> *Subtype:* Person
>> *Attribute:* Name
>> *Attribute:* Date

Attribute: Biography
Subtype: Corporate Body
　　Attribute: Name
　　Attribute: Date
　　Attribute: Place
　　Attribute: Administrative history

5.2.3 Producer

An Agent responsible for the existence of the physical or electronic form in which an Item is realised.
Note: The term as used here encompasses agents with a variety of roles. It may encompass publisher, scribe, printer, binder, distributor &c.
　Subtype: Person
　　Attribute: Name
　　Attribute: Date
　　Attribute: Biography
　Subtype: Corporate Body
　　Attribute: Name
　　Attribute: Date
　　Attribute: Place
　　Attribute: Logo
　　Attribute: Standard number
　　Attribute: Administrative history

5.2.4 Collector

An Agent who gathers Items together.
　Subtype: Person

Attribute: Name
Attribute: Date
Attribute: Biography
Subtype: Corporate Body
　Attribute: Name
　Attribute: Date
　Attribute: Place
　Attribute: Administrative history

5.2.5 Owner

An Agent who has legal possession of a Collection.
Note: The transfer of ownership carries with it the transfer of any Owner's rights, so the persistence of rights is not usually an issue. However, Previous Owners have had possession of the objects themselves, and are frequently a legitimate focus of interest. It is also possible that conditions of sale may circumscribe the rights of a subsequent Owner.
　Subtype: Person
　　Attribute: Name
　　Attribute: Date
　　Attribute: Biography
　Subtype: Corporate Body
　　Attribute: Name
　　Attribute: Date
　　Attribute: Place
　　Attribute: Logo
　　Attribute: Administrative history

5.2.6 Administrator

An Agent who has responsibility for the physical or electronic environment in which a Collection is held.
Note: This is not necessarily the Owner: e.g. a Collection may be on deposit or loan.

> *Subtype:* Person
>> *Attribute:* Name
>> *Attribute:* Date
> *Subtype:* Corporate Body
>> *Attribute:* Name
>> *Attribute:* Date
>> *Attribute:* Place
>> *Attribute:* Logo
> **Note**: Should Affiliation should be considered e.g. HEFCE?

5.3 Entities: Indirect Agents

5.3.1 Each of the above Agents has a direct link with one of the objects of concern. There are also other agents whose involvement is more indirect.

5.3.2 Both Creators and Producers have rights which persist beyond the actual creation and publication of a work, and which remain relevant for those wishing to use the work. These rights may outlast the Creator or Publisher, or may be transferred by them to a third party. In either case an new Agent is involved who is relevant to the description of a Collection.

5.3.3 Creator's Assignee

An Agent holding rights originally vested in the Creator.

> *Subtype:* Person
>> *Attribute:* Name
>> *Attribute:* Date
>> *Attribute:* Biography
> *Subtype:* Corporate Body
>> *Attribute:* Name
>> *Attribute:* Date
>> *Attribute:* Place
>> *Attribute:* Logo
>> *Attribute:* Administrative history

5.3.4 Producer's Assignee

An Agent holding rights originally vested in the Producer.

> *Subtype:* Person
>> *Attribute:* Name
>> *Attribute:* Date
>> *Attribute:* Biography
> *Subtype:* Corporate Body
>> *Attribute:* Name
>> *Attribute:* Date
>> *Attribute:* Place
>> *Attribute:* Logo
>> *Attribute:* Administrative history

5.4 Relationships

5.4.1 Each of the primary Agents is linked to some aspect of the objects in question. In many cases the nature of the relationship is itself complex, and there are attributes which relate to the link itself rather than to either of the entities linked. Note that the following paragraphs deal with relationships between the entities involved in Collection Description; the relationship between Collection Descriptions themselves is a different matter discussed in section 7.

5.4.2 In the accompanying chart, links are to be read as far as possible from the top and from the left. For example, the Content (top) *Is-Embodied-In* the Item (bottom); the Creator (left) *Creates* the Content (right). In each case the reciprocal statement of the link is implied, e.g. the Content *Is-Created-By* the Creator. The names given to the Relationships are self-explanatory and are not redefined separately; they are, however, given additional glosses where they carry meanings beyond the actual words used. Where the same words would otherwise describe two relationships occurring in the chart, an initial code letter has been added to distinguish the two instances.

5.4.3 Holds-Rights-Of

A Relationship between an Agent and an Indirect Agent embodying the rights previously linking the Agent to an Object.
Note: There are two instances of this Relationship in the Collection model, embodying Creator's and Producer's rights.

> *Attribute:* Nature of transfer [Bequest, Sale &c]
> *Attribute:* Terms of transfer
> *Attribute:* Date

5.4.4-Sells-To

A Relationship between Agents embodying transfer of ownership of an Object or Objects.
Note: Sells-To is merely the name given to the relationship and may represent any of the modes of transfer of ownership e.g. donation or bequest. A Previous Owner may impose certain constraints on the present Owner
> *Attribute:* Nature of transaction
> *Attribute:* Terms of transaction
> *Attribute:* Rights.....
> *Attribute:* Date

5.4.5 Contracts-With

A Relationship between Creator and Producer empowering the Producer to produce an Item or Items embodying Content created by the Creator.
Note: The name of the relationship implies a commercial contract as between an author and a commercial publisher, but in this model encompasses any arrangement linking the two Agents as defined.
> *Attribute:* Terms of contract
> *Attribute:* Date

5.4.6 Delegates-To

A Relationship between an Owner and an Administrator empowering the Administrator to incorporate a Collection in a Location
Note: The Delegates-To relationship may include certain rights or constraints which the owner imposes on the Administrator.
> *Attribute:* Terms of delegation
> *Attribute:* Date

5.4.7 Creates

A Relationship between a Creator and Content, specifying the manner of creation and the rights associated with the act of creation.
Note: The Creates relationship can take many forms: writing, composing, performing, editing &c. Each of them may have some sort of right associated with it.
> *Subtype:* Authoring
> > *Attribute:* Copyright
> > *Attribute:* Moral right
> *Subtype:* Singing
> > *Attribute:* Performance right
> *Subtype:* Inventing
> > *Attribute:* Patent and registered design right
> &c.

Note: (1) A fuller list is given in Appendix B, derived from the Library of Congress list of relator codes. (2) In UK legislation a moral right separate from copyright is recognised for authors in the broad sense.

5.4.8 Produces

A Relationship between a Producer and an Item or Collection-Description specifying the manner of production and the rights associated with the act of production.
Note: Like Producer, the Produces relationship should be interpreted in the broad sense to include printing, distribution &c. Again, rights are involved.
> *Subtype:* Publishing
> > *Attribute:* Copyright
> *Subtype:* Printing
> *Subtype:* Writing
> *Subtype:* Pressing
> *Subtype:* Distributing
> &c.

5.4.9 Collects

A Relationship between a Collector and a Collection specifying the manner in which a Collector causes the Collection to accrue.
> *Attribute:* Legal status
> *Attribute:* Accrual policy (closed; passive; active; partial/selective)

5.4.10 Owns

A Relationship between an Owner and a Collection or Collection-Description specifying the rights associated with ownership.

Note: The Owner expresses ownership by exercising such rights as controlling access to and reproduction of the material.

> *Attribute:* Legal status
> *Attribute:* Access control [Allowed users, Charges &c]
> *Attribute:* Maintenance [Conservation action &c]

5.4.11 Administers

A Relationship between an Administrator and a Location specifying the manner in which the Administrator administer the Location.
Note: The Administrator Administers the Location by opening and closing it; and by admitting or not admitting classes of user to the Location, on certain terms. In many – most – cases the Administrator will be the same Agent as the Owner; or may be exercising rights of the Owner that have been Delegated-To the Administrator. (For example, there may be distinct charges for access to the Location and to the Collection.)

> *Attribute:* Access conditions [Hours of access, classes of permitted user, &c]

5.4.12 Is-Embodied-In

A Relationship between Content or Content-Component and Item or Collection-Description, specifying the manner in which the Content or Component is realised in the Item.
Note: Intellectual Content is only accessible to people other than its Creator by finding embodiment in some physically perceptible form. The attributes of the method of embodiment depend on the nature of the document: printing, pressing, &c.

Natural objects (e.g rocks) which may be part of a Collection do not have intellectual content,therefore the Relationship may be empty.

> *Attribute:* Infixion method (print; photolithography; handwriting; recording method; groove type; track type)
> *Attribute:* Infixion characteristics (typeface; script; colour characteristics; polarity; aspect ratio; reduction ratio; playing/projection speed; recording density; sectoring)

Note: See *Analytical Model of AACR, Part I*, Tab 11 on infixion.

5.4.13 Is-Gathered-Into

A Relationship between Items or Item-Components and a Collection specifying the manner in which Items or Components are or have been gathered into the Collection
Note: The essence of a Collection is that diverse material has been gathered together.

> *Attribute:* Accrual method (legal basis (purchase, deposit &c.))
> *Attribute:* Accrual periodicity (closed, irregular, periodic
> *Attribute:* Density
> **Note:** Density refers to whether the collection is Compact, i.e. the Items in it are gathered, for practical purposes, into a single block of material; or Diffuse, i.e. the Items in it are physically scattered through a larger Collection (which may or may not have its own Collection-Description)
> *Attribute:* Identifier
> **Note:** The Identifier relating an Item to a Collection will usually be a call number/shelfmark. Distinguish from the Item Identifier which may be an ISBN or equivalent.

5.4.14 Is-Located-In

A Relationship between a Collection or Collection-Description and a Location specifying the manner and limits of the deposition of a Collection in its Location.
Note: 'Locating' may be physical or electronic.
    Attribute: Dates of deposit

5.4.15 Has-Part

A Relationship between Content or an Item and an instance of a Component of the Content or Item.
    Attribute: Relation identifier
    Note: The Relation identifier enables the placement of the Component relative to the whole Content or Item to be determined, e.g. 'Book 1 of 3'

5.4.16 Is-Described-By

A Relationship between a Collection and/or its Items on the one hand and a Collection-Description on the other.
Note: The Is-Described-By Relationship stands outside the Entities and Relationships which comprise Collections; it links the realms of data and metadata. The metadata entities are described below.
     Attribute: Currency

5.5 Entities: Collection-Description

The Class of Entities 'Collection-Description' consists of entities providing information about Entities of the Class 'Objects'.

5.5.1 Entity: Unitary Finding-Aid

A Collection-Description which consists only of information about the Collection as a whole and does not provide information about the individual Items within it.
    Attribute: Cataloguing code
    Attribute: Entities and Relationships present
    Note: This attribute will have as its data elements the list of Entities and Relationships present in the Collection model which are recorded in this Collection-Description.

5.5.2 Entity: Hierarchic Finding-Aid

A Collection-Description which consists of information about the Collection as a whole, together with information about the individual Items within it and their Content, including contextual information about the relation of the Items and their Content to the Collection as a whole.
    Attribute: Cataloguing code
    Attribute: Entities and Relationships present
    Note: This type of Collection-Description is most often associated with archival collections where contextual information is necessary to the understanding of the Items.

5.5.3 Entity: Analytic Finding-Aid

A Collection-Description which consists of information about the individual Items within it and their Content.

> *Attribute:* Cataloguing code
> *Attribute:* Entities and Relationships present
> **Note:** This type of Collection-Description is typical of a

library catalogue.

5.5.4 Entity: Indexing Finding-Aid

A Collection-Description which consists of information derived from the individual Items within it.

> *Attribute:* Cataloguing code
> *Attribute:* Entities and Relationships present

Note: See the next section, section 6, Collection Description, for further discussion of these definitions.

6. Collection-Description

6.1 Collection-Description can take any of the forms listed in Section 5.5. The subtype of Collection-Description will determine the entry-point at which the Collection-description is linked to the Entity-Relationship diagram for the Collection, thereby determining how the Collection model will be traversed, and which of those Attributes of the Entities and Relationships encountered will be incorporated into the Collection-Description.

6.2 In three of the four identified types of Collection-Description the information it conveys is analytic: that is, the information is held in discrete packets (e.g catalogue records) which, although they may be brought together and presented as a result of a search, or may be organised in a particular sequence (e.g. by author's name), are largely independent of each other.

6.3 Two qualifications have to be made to the paragraph above. First, a Collection-Description may have some overall structure which reduces the autonomy of its constituent elements – i.e. it may be necessary to know the placement of a catalogue record within the structure of the catalogue – its context -- in order to interpret the record correctly. This is always true to some extent, and the participants in the Toronto conference on the principles and development of AACR stressed the weaknesses in online catalogues resulting from the loss for such contextual information (for example, the ordering of results sets is often effectively arbitrary). It is particularly true for the established practices in cataloguing archival collections (see the rules for multilevel description in ISAD(G)).

6.4 Second, with internet resources the distinctions may become blurred. Take, for example, the existence of a site for the works of Kipling on the World Wide Web. Viewing the site as a whole, it may be said to be a collection of entities or derived representations of entities. However, if much the same list of links can be retrieved by a search on (say) Yahoo, does this make Yahoo a Collection in our definition instead of a Collection-Description? This study takes the view that ownership, administration and location are relevant to the definition of a collection. The fact that a catalogue can now be directly linked to the entities catalogued – that the searcher can move seamlessly from finding and identifying to selecting and

obtaining – need not mean that the constituent elements of those processes have changed.

6.5 In the diagrams in section 4.2, the different types of Collection-Description are shown as distinctly and discretely arising from different elements of the Collection diagram. This is a simplification, but one which does not materially affect the model, as all of the elements of the Collection model are available to the Collection-Description model.

6.6 A Unitary Finding-Aid takes as its basis the information about the Collection as a whole – it makes no attempt to capture information about individual records except in so far as it is necessary to provide aggregate information (e.g. on limiting dates, or on the number of Items it contains).

6.7 An Analytic Finding-Aid lists the individual records comprising information about the intellectual Content and the Items in which it is realised. There may, in the individual records, be information about Collections, and the Finding-Aid may be searchable from that aspect, but that is not its focus. A library catalogue is typically an Analytic Finding-Aid.

6.8 An archival collection is more often described by a Hierarchic Finding-Aid, in which the individual Items and their Content are described, but firmly grounded within the overall arrangement of the Collection, e.g. grouping together all the letters, account books &c. in an ordered sequence or sequences. The Items are often not uniquely identifiable when considered in isolation, so the context of the Collection is an essential element in compiling the Collection-Description.

6.9 An Indexing Finding-Aid is characterised here as consisting of information derived from Items, by implication regardless of their Content. By this is meant that an Indexing Finding-Aid – such as a robotic search engine – will index the words in a document (or catalogue record) regardless of their context and without trying to identify the discrete elements of Content contained therein. The effects of this may be mitigated by the use of metadata tags in web documents, but in so far as such tags are used by the engine, it is creating an Analytic Finding-Aid (which may or may not be combined with the Indexing Finding-Aid). An online Analytic Finding-Aid may incorporate a keyword index which is, in effect, an Indexing Finding-Aid in this sense of the term. At the other end of the technological scale, a printed Calendar of a Collection may have its own printed Indexing Finding-Aid which lists -- out of context -- the names, places &c occurring in the Collection.

7. External Relationships

7.1 Because it is a model of a single instance of a Collection, the model of Collection Description does not explicitly map external relationships. Such relationships are between instances of the model, and are not part of the internal structure of the model itself. They may, moreover, operate both at the Collection level and at the Collection-Description level. Relevant external relationships are:

Has-Version (including format distinctions; the distinctions made in CCSDS *Reference Model for an Open Archival Information System (OAIS)* section 5.1.3 p. 5-4 may appropriately be used as instances of this class of Relationship)

Has-Part (relating parts to wholes; discussion has identified the desirability of distinguishing Diffuse)

Has-Complement (where a formerly unified Collection has been split into separate parts, possibly in separate institutions)

Has-Association (where another Collection is relevant but the relationship is not described by one of the above list. Typically relates to another Collection associated with one or more of the Agents of the Collection)

Has-Publication (where the collection is the basis of a published scholarly work). The associated Publication is unlikely to be present in a Collection Description resource so this External Relationship will not be reciprocal

Is-Described-By (relating a Collection or Collection-Description to a Collection-Description)

7.2 The relationship between the Collection-Description model and the external relationships is that of catalogue record to catalogue. The rules for constructing catalogue records do not of themselves dictate the form of catalogue, although they carry strong implications for it.

7.3 Construction of a catalogue involving multiple records and cataloguing tools requires knowledge of the intended purposes of the catalogue. At the same time it is now very often the case that the same record or records will be found and used in several different catalogues, each with their own purposes. In such a context, although it is useful and necessary to ask what searches one might envisage in using a particular Collection-Description tool, there is also benefit to be gained from keeping the underlying model as comprehensive, as generic and as open as possible, so that a particular implementation of the model has a better chance of being relevant to and usable in another Collection-Description tool.

8. Users

8.1 Many traditional catalogues and indexes have not incorporated information about access conditions fully, or even at all. It is not usual, for example, to find details of a library's opening hours in the catalogue records for its holdings. In an online environment, however, the issues of rights of access and use become more prominent. Instead of selling a hard-copy item a publisher may want to recover costs by retaining the material and selling access to it as a database.

8.2 This model, therefore, tries to clarify the points at which rights and conditions of access and use become operable. The way in which those conditions are realised may vary, and may involve different classes of condition: restrictions relating to time, place, class of person, purpose of use &c may need to be elaborated. A knowledge of the totality of conditions of access and use is required in a database which attempts to act as a bridge linking collections and their users.

Appendix A: Mapping of concepts in relevant work

For each Entity and Relationship in the model other than Components, comparable elements in each of the various schemes are listed. As the attributes of the three Component Entities are the same as those of their related wholes, they are not compared explicitly in the following list. If a scheme is not listed under any of the headings, it implies that the scheme has no explicit element which can be mapped to this heading.

A..1 Content
    DC:
        Title
        Description
        Subject
        Date
        Identifier
        Language
        Coverage
    UKOLN:
        Title
        Description
        Subject
        Date
        Type
        Identifier
        Source
        Language
        Coverage
        Notes
    ROADS:
        Title
        Keywords
        Subject descriptor
        Subject descriptor scheme
        Short title
        Alternative title
        Language
        Category
    MODELS:
        Title
        Subject-Heading
        Standard-Number
    FRBR:
        Work
        Expression
        Concept
        Object
        Event
        Place [subject]
    ISAD:
        Title
        Dates of creation
        Scope & content/ Abstract
        Language of material
    MASTER:
        msContents

A.2 Item
    DC:
        Title
        Format
        Date
        Identifier
        Coverage
    UKOLN:
        Title
        Date
        Identifier
        Source
        Coverage
        Notes
    ROADS:
        Title
        Short title
        Alternative title
        Handle
        Format
        Description
        ISBN
        ISSN
        Category

Comments
Character-Set
Citation
Size
MODELS:
Control-Number-Local
Date-publication
FRBR:
Manifestation
Item
ISAD:
Extent
Physical characteristics
MASTER:
physDesc

A.3 Collection
DC:
Title
Description
Subject
Date
Type
Identifier
Source
Language
Coverage
Format
UKOLN:
Title

Description
Subject
Date
Type
Identifier
Source
Language
Coverage
Notes
ROADS:
Title
Description
Handle
Keywords
Subject descriptor
Subject descriptor scheme
Short title
Alternative title
Language
Source
Format
Creation-Date
Category
Comments
Citation
Size
MODELS:
Title
Subject-Heading
Standard-Number

Control-Number-Local
Date-publication
ISAD:
Extent
Dates of accumulation
Scope & content/ Abstract
System of arrangement
MASTER:
msContents
physDesc

A.4 Location
DC:
Identifier
UKOLN:
Identifier
ROADS:
URL
Keyword-Places
ISAD:
Reference code
MASTER:
msIdentifier

A.5. Creator
DC:
Creator
Contributor
ROADS:

Author
Keyword-Organizations
Keyword-Names
MODELS:
Author-name
FRBR:
Person
P-name
P-date &c
Corporate body
C-name
C-place
C-date &c
ISAD:
Name of Creator
Admin/ Biographical
history
MASTER:
author

A.6 Producer
DC:
UKOLN:
Logo
ROADS:
Publisher
Keyword-Organizations
Keyword-Names
MODELS:
Standard Number

FRBR:
[M] Place of publication
[M] Publisher/ distributor
[M] Fabricator/
manufacturer

A.7 Collector
DC:
Creator[?]
FRBR:
Person
P-name
P-date &c
Corporate body
C-name
C-place
C-date &c
ISAD:
Name of Creator
Admin/ Biographical
history

A.8 Owner
UKOLN:
Owner
Logo
ROADS:
Owner
Sponsor
ISAD:

Immediate source of
acquisition
MASTER:
history

A.9 Administrator
UKOLN:
Administrator
ROADS:
Admin

A.10 Creator's Estate
[Not in any of the schemes]

A.11 Publisher's Estate
[Not in any of the schemes]

A.12 Holds-Rights-Of
DC:
Rights
UKOLN:
Rights
Use constraints
ROADS:
Copyright
ISAD:
Copyright / Conditions
governing reproduction

A.13 Sells-To
    ISAD:
        Custodial history
    UKOLN:
        Charging policy
    FRBR:
        [I] Provenance

A.14 Contracts-With
    [Not in any of the schemes]

A.15 Delegates-To
[Not in any of the schemes]

A.16 Creates
    DC:
        Rights
    UKOLN:
        Rights
    ROADS:
        Copyright
    ISAD:
        Copyright / Conditions
governing reproduction

A.17 Produces
    DC:
        Rights
    UKOLN:
        Rights

        Use constraints
    ROADS:
        Copyright
        Publication-Status:
    ISAD:
        Copyright / Conditions
governing reproduction

A.18 Collects
    ISAD:
        Appraisal, destruction,
        scheduling info
        Legal status
        Accruals

A.19 Owns
    DC:
        Rights
    UKOLN:
        Rights
        Use constraints
    ISAD:
        Appraisal, destruction,
        scheduling info
        Legal status
        Accruals

A.20 Administers
    UKOLN:
        Use constraints

        Access policy
    ROADS:
        Access-Policy
        Access-Times
        Charging-Policy
        Registration
        Requirements
    FRBR:
        [I]Access restrictions
    ISAD:
        Access conditions
    MASTER:
        availability

A.21 Is-Embodied-In
[Not in any of the schemes]

A.22 Is-Gathered-Into
    ISAD:
        Reference codes

A.23 Is-Located-In
[Not in any of the schemes]

A.24 Is-Described-By
    DC:
        Source [of derived item]
        Relation

A.25 Unitary Finding-Aid
> ROADS:
>> Destination
>> To-be-reviewed date
>> Record-last-verified email
>> Record-last-verified date
>> Record-last-modified email
>> Record-last-modified date
>> Record-created email
>> Record-created date
>> Checked-by-Name
>> Checked-by-Date
> MASTER:
>> recordHist
>> bibl

A.26 Hierarchic Finding-Aid
> ISAD:
>> Level of desc
>> Finding aids

A.27 Analytic Finding-Aid
> UKOLN:
>> Collection.Catalogue
> FRBR:
>> Aggregate
>> Component

A.28 Indexing Finding-Aid
[Not in any of the schemes]

A.29 Other, not reflected in the model
> ROADS:
>> Discussion
>> Source
>> Last-Revision-Date
>> Template-Type
>> Version
> FRBR:
>> [I] Exhibition history
> ISAD:
>> Location of originals
>> Existence of copies
>> Related units of description
>> Associated material
>> Publication note
>> [**NB** All the above are noted as External Relationships]
>> Note
> MASTER:
>> surrogates

Appendix B: Types of Agent-Object Relationship (based on the *US MARC code list for relators, sources, description conventions).*

Creates **Relationship** *(5.4.7)*

Performing:

Acting
Dancing
Playing [an instrument]
Performing
Singing
    Attribute:
      Performance right

Constructing:

Acting as architect
Engineering
Inventing
    Attribute: Patent and
    registered design right

Authoring:

Adapting
Annotating
Drawing
Authoring
Censoring

Choreographing
Collaborating
Commentating
Composing
Corresponding
Costume designing
Curating [an exhibition]
Dedicating
Designing [a stage/film production]
Directing
Editing
Film editing
Hosting
Illuminating
Illustrating
Inscribing
Interviewing
Moderating
Narrating
Originating
Photographing
Programming
Redacting
Reviewing
Scientific advising
Sculpting

Signing
Songwriting
Speaking
Advising on a Thesis
Translating
Writing a Libretto
Writing a Scenario
    Attribute: Copyright
    Attribute: Moral right

Produces **Relationship** *(5.4.8)*

Arranging
Binding
Binding designing
Book designing
Book producing
Bookjacket designing
Bookplate designing
Bookselling
Calligraphing
Cartographing
Collotyping
Compiling
Compositing
Conducting

Correcting
Distributing
Electrotyping
Engraving
Etching
Funding
Licensing
Lithographing
Metal-engraving
Papermaking
Platemaking
Printing
Printing of plates
Producing
Proofreading
Publishing
Recording
Rubricating
Sponsoring
Stereotyping
Transcribing
Type designing
Typographing
Wood-engraving
    Attribute: Copyright

Appendix C: References

A Complete Collectuion of Poems by Rudyard Kipling
(http://www.rit.edu/~exb1874/mine/kipling/kipling_ind.html)

Dublin Core Metadata Element Set: Reference Description
(http://purl.org/dc/dc-elements.htm)

Functional Requirements for Bibliographic Records
(http://www.ifla.org/VII/s13/frbr/frbr.pdf)

ISAD(G): General International Standard Archival Description
(http://data1.archives.ca/ica/cds/ isad(g)e.html)

The Logical Structure of the Anglo-American Cataloguing Rules,
drafted for the Joint Steering Committee for Revision of AACR -
Part 1 by Tom Delsey (National Library of Canada) with assistance
from Beth Dulabahn (Library of Congress); Michael Heaney
(Oxford University); Jean Hirons (Library of Congress) - August
1998; *Part II* by Tom Delsey (National Library of Canada) with
assistance from Beth Dulabahn (Library of Congress); Michael
Heaney (Oxford University) - January 1999. (http://www.nlc-
bnc.ca/jsc/)

*Metadata: Collection Level Description: Collection description
working group: Work in progress, 12 October 1998* / Edited by:
Andy Powell, UKOLN (http://www.ukoln.ac.uk/metadata/cld/ wg-
report/)

MODELS Profile Interoperability Sub-Set
(http://www.m25lib.ac.uk/M25link/modelssubset.html)

The principles and future of AACR / Jean Weihs, editor. (Ottawa:
London Chicago: Canadian Library Association; Library
Association Publishing; American Library Association, 1998).
(Proceedings of the International Conference on the Principles and
Future Development of AACR, Toronto, Ontario, Canada, October
23-25, 1997).

Reference Model for an Open Archival Information system (OAIS)
Consultative Committee for Space Data Systems, May 1999
(available from http://www.ccsds.org/RP9905/RP9905.html)

ROADS Cataloguing Guidelines: *Revised version (v. 1.0)* / by
Michael Day (http://www.ukoln.ac.uk/metadata/roads
/cataloguing/cataloguing-rules.html)

*Towards a European Standard for Manuscript Description:the
MASTER project /* Lou Burnard and Peter Robinson
(http://www.cta.dmu.ac.uk/projects/master/)

US MARC code list for relators, sources, description conventions /
Library of Congress. Network Development and MARC Standards
Office. Web edition, 21 January 1999.
(http://lcweb.loc.gov/marc/relators/)

Yahoo (http://www.yahoo.com)