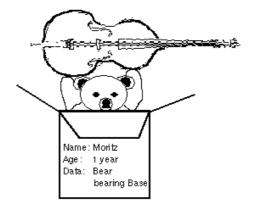
# The Metadata Bear

Or: Bearing the weight of accessibility



Carola Boehm, (Database Consultant, Performing Arts Data Service), Department of Music, University of Glasgow, carola@music.gla.ac.uk

#### 1. Introduction and Background

The Performing Arts Data Service (PADS)<sup>1</sup> is one of a syndicate of five Service Providers appointed by the Arts and Humanities Data Service (AHDS)<sup>2</sup>, funded by the Joint Information Systems Committee (JISC)<sup>3</sup> of the UK's Higher Education Funding Councils, and is based at the University of Glasgow. The AHDS's mission is to co-ordinate access to, and facilitate the creation and use of, electronic resources in the arts and humanities by offering a range of services. It will encourage scholarly use of its collections and make information about them available through an on-line catalogue.

The AHDS provides a single gateway for arts and humanities scholars wishing to search for data-sets across various discipline areas. Other service providers include: the History Data Service (HDS)<sup>4</sup>, The Archaeology Data Service (ADS)<sup>5</sup>, the Oxford Text Archive (OTA)<sup>6</sup> and the Visual Arts Data Service (VADS)<sup>7</sup>. The service providers' databases interoperate with other databases within the AHDS and beyond via Z39.50<sup>8</sup>, and searching is available via the Web. In order to achieve meaningful search results, data from all the service providers is indexed with Dublin Core metadata.

The Performing Arts Data Service's role within this framework is to support research and teaching in UK Higher Education by collecting and promoting the use of digital data relating to the performing arts: music, film, broadcast arts, theatre and dance. The PADS differs from the other service providers in that it has a particular concern with data consisting of and representing time-based media. This paper will present the specific issues of implementing a digital library service in the performing arts, emphasising on metadata management in particular.

#### 2. Data and Metadata in the Performing Arts

In a time where the internet becomes the platform, the browsers become the operating system, and applications become services a digital library project set in the performing arts has to define new methods of storing and distributing time-based data to be able to serve quality and quantity information across wide area networks. It also requires solutions of the more philosophical areas of research into how interfaces have to be set up and how information should be represented in order for users to handle vast amounts of data as intuitively and as user friendly as possible. In order to cope with the information increase and user capacity expected information management services will always have to deal with the "three I's": Information Structure, Information Representation and Information Access.

<sup>&</sup>lt;sup>1</sup> http://www.pads.ahds.ac.uk, 1999-06-11.

<sup>&</sup>lt;sup>2</sup> http://www.ahds.ac.uk/ , 1999-06-18.

<sup>&</sup>lt;sup>3</sup> http://www.jisc.ac.uk , 1999-06-18.

<sup>&</sup>lt;sup>4</sup> http://www.hds.ac.uk/ 1999-05-08.

<sup>&</sup>lt;sup>5</sup> http://www.ads.ac.uk/ 1999-06-11.

<sup>&</sup>lt;sup>6</sup> http://www.ota.ac.uk/ 1999-06-11.

<sup>&</sup>lt;sup>7</sup> http://www.vads.ac.uk/ 1999-06-11.

<sup>&</sup>lt;sup>8</sup> http://lcweb.loc.gov/z3950 , 1999-06-18.

| <ul> <li>primary resource: representations of the work itself</li> <li>recordings of the work         <ul> <li>recording of a first performance of the work</li> <li>recording of a performance with the composer</li> <li>recordings of other performances</li> </ul> </li> <li>facsimiles of the work         <ul> <li>facsimile of score written by composer</li> <li>facsimile of first publication of the work</li> </ul> </li> <li>video as the work</li> <li>complex-music-data data as the work(e.g. MAX)</li> <li>scripts/texts as the music resource (e.g. action/text based scores)</li> <li>graphics as the music resource (e.g. graphic scores)</li> <li>printed score</li> </ul> | <ul> <li>secondary resources: resources about the work         <ul> <li>recordings of                 <ul> <li>interviews                     <ul> <li>features</li> </ul> <li>facsimiles of letters/notes containing information<br/>about the work</li> <ul> <li>video as secondary information (composer, performance, etc)</li> <li>image as secondary information (composer, performance, etc.)</li> <li>reviews</li> <ul> <li>analytical texts, diagrams and images</li> </ul> </ul></li> <ul> <li>images</li></ul></ul></li></ul></li></ul> |
|--|--|
|--|--|

The information structure of data relating to the Performing Arts is by nature diverse: from text based, to visuals/images, to the intrinsically time-based media. Resources of Music, film and video, theatre and performance art, broadcasting arts and dance have their special needs and requirements. Any information management system dealing with this range of material must be able to store complex and composite data, cope with a multitude of single documents, and offer intelligent, user-friendly but controlled access over wide area networks.

The common characteristic of all this data is that it is either inherently time-based or that it is a part of a time-based entity or that it describes/displays a time-based entity. This major characteristic of being time-based demands for specialist solutions for handling the information representation for and access to these resources. One of the vital questions need to be addressed is: How is metadata able to support the "time-based-ness" in order to facilitate intelligent access and delivery for the performing arts user community?

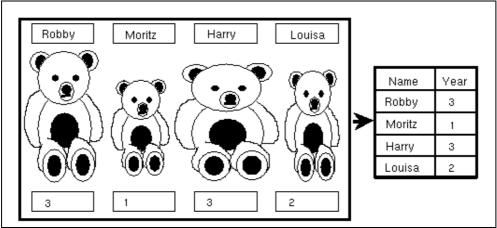


Fig.2 Data, Metadata and Databases

Metadata is in general defined as "Data about data". It is probably correct to say that the library community is genuinely the community which has had the advantage of the longest experience of dealing with metadata in the form of catalogues in order to find data in form of physical textual resources. But is probably wrong to assume, that this bibliographic metadata is sufficient for discovery, access and management digital resources.

General requirements of metadata for digital resources would include amongst others:

- metadata for search and retrieval (in one database, across distributed databases (and objects), across disciplines)
- metadata for user system management processes such as access and security, licensing, delivery mechanisms
- metadata for mapping knowledge domains intelligently and archiving cultural heritage
- metadata for searching and managing the content

<sup>&</sup>lt;sup>9</sup> See http://www.music.gla.ac.uk/HTMLFolder/Research/smatbam-private/categories.html, 1999-06-11 for a discussion of this issue.

Thus, if we look at existing standards and protocols for metadata we have a multitude of different technologies which might have to be supported:

- Metadata of bibliographic nature (AACR, MARC, Bib1, etc))
- Metadata and Interoperability protocols (Z39.50, ZORBA, etc)
- Meta-Metadata (Dublin Core, RDF, etc)
- Metadata and DBMS (SQL, etc)
- Metadata and Object-oriented Technologies (OQL, MVC)
- Metadata and Content Searching (MPEG7, SMPTE)

Metadata thus becomes the major "access technology" for any user. The metadata management of digital resources will have to be considered most carefully in any kind of digital library service. This is what I call the "Metadata Bear", the technology "bearing the weight of accessibility".

For a digital Library Service, such as the PADS, whose remit is to "collect and promote the use of digital data resources to support research and teaching in the performing arts"<sup>10</sup> the digital data resources include a) catalogues, b) digital resources themselves, d) digital representations of the resources, and e) digital versions of works. The "collecting and promoting" of these resources becomes in a major way a responsibility of "Providing Access through Metadata".

## 3. A Digital Library Service – Promoting Access through Metadata

When considering what functionality a "Metadata Management System" should encompass in order to support a digital library service, three aspects should be precisely defined and specified: a) the users, b) the data and c) the access.

## 3.1. The users

In a digital resource collection service there are different types of users. This has to be taken into account when implementing secure access with rights management, i.e. the maintaining and administering of access rights and licenses. Time-based media collection owners, such as museums, music publishers and labels, will only be willing do offer the use and the publishing of certain collections if restricted access can be guaranteed, and sometimes only if the management of the rights is transparent and administered by the collection owners themselves. The dependency of these bodies on their collection as a means of income, as well as the maintenance of existing licenses and access rights of their collection, will be a important issue when discussing collection input.

| USERS  | RIGHTS MANAGEMENT  |
|--|--|
| Providers  | Providers  |
| Developers   | <ul> <li>full access, configuration and programming rights</li> </ul>          |
| Collection Officers  | full access –  |
| Project Managers   | full access  |
| Close associates   | Close associates   |
| <ul> <li>Content owners and providers</li> </ul>                   | <ul> <li>full access and management rights to parts of a collection</li> </ul> |
| End Users  | End Users  |
| <ul> <li>user groups with rights to certain collections</li> </ul> | <ul> <li>read access to part of a collection</li> </ul>                        |
| anonymous users  | <ul> <li>read access to public collections</li> </ul>                          |

Fig. 4 Possible user levels and rights management

This of course implies a system of (at least) hierarchical management of access rights of users and groups, thus the management of access/licensing rights as possible metadata of a resource or the collection in which it is located.

#### 3.2. The data

It is desirable that a collection is able to be expanded by collections of other service providers holding resources in the same field but at the same time maintaining a "one-stop shop" in accessing timebased media resources. This distributed resource environment allows the option of other collection holders keep and maintain their collection physically in their own repository, while access is handled by a central access point.<sup>11</sup>

<sup>&</sup>lt;sup>10</sup> Performing Arts Data Service, PADS Collections Policy,

http://www.pads.ahds.ac.uk/PadsUserServicesCollection, 1999-06-18.

<sup>&</sup>lt;sup>11</sup> This distribution of information has also implications reagarding copyright. I.e. Institutions holding copyright of material will likeky want to hold their collection physically on their servers and still be able to offers single user interfaces across remote collections. The National Preservation Office of the National Library of Australia has

A performing arts resource collection encompasses a wide range of different disciplines, starting with the disciplines of music and film and stretching further toward dance, theatre and the broadcasting arts.<sup>12</sup> The resources as a whole can be characterised as a) being made out of different types of data, b) containing differing complexities of data, c) possessing different relationships.

#### 3.2.1 Different types of data

As with all multimedia related systems, all the "usual" data types are involved from sound, video, text, image and binaries. Storing them in a certain way provides us with a more complex entity of data types: html, sgml, mpeg, wav, gif, jpeg, java, etc. It is certain that these data formats will evolve further in number and content. The use of different formats in a system should therefore be a means but not a solution. In other words, to minimise the danger of storing data in standards that might not be supported in the future, much thought should go into separating the content of a resource from its presentations. To be able to store a resource in the highest quality possible, combined with the ability to convert it into formats suitable for a certain purpose, or added formats in the future, is to provide an open and flexible system with maximum compatibility in the long term.

#### 3.2.2 Differing complexity of data

Whereas video and images might be stored largely as single binary data-objects, music, theatre and the broadcasting arts could involve the storing and accessing of highly structured data, presenting complex objects or 'composite objects'.<sup>13</sup> In some cases, it might be hard to distinguish which is the real, the original resource, and which is a composite part of it. If one accepts the fact that the content of a resource might be of complex or composite nature, then the step towards devising a way to store it as such is not far. Technologies are needed that offer the ability to depict, represent, access, store and manipulate complex structures in their complex "Gestalt". A broadcasting feature, as one resource, might encompass video data, sound data, and text data and still be one work of art.

We should accept the fact that our future data might not remain in its binary form and much of our present resources have never been in the "Gestalt" of one entity. Java Applets, Webobjects and other distributed object environments are already being used by artists to create works made out of many components and having many facades. Also the existing resources, which have been traditionally stored as metadata in catalogues, while their real content is being stored as artefacts in shelves, cassettes, or discs, are often not just one entity. In trying to devise resource systems of the next decade, it would be illogical to diminish the resources and their "real-life" manifestation by disregarding their composite character. It was clear for achieving the above requirements, normal library catalogues and conventional relational database management systems would not be sufficient. Object-oriented or at least object-based information system technologies would have to be employed.

#### 3.2.3 Different relationships

Assuming that we have objects stored in a persistent way, the access and search results are influenced by the context these objects are in. Not only that, our minds create sense out of contexts of informations (sic). Data without context may be hard to interpret or to map intelligently. The mapping of content and context into a digital world means defining and storing different kinds of relationships between objects.<sup>14</sup> Relationships can be of numerous variety. For example, five basic relationships widely used in information systems are:

• *Inclusion* - one object is included in another object (e.g. a file in a folder, a certain sound used in a composition, a note in a bar)

made this "Distributed responsibility" one of its Statements of Principles of Preservation of and Long-Term Access to Australian Digital Objects. See http://www.nla.gov.au/nla/staffpaper/preserve.html, 1998-07-01.

<sup>&</sup>lt;sup>12</sup> See Categories of time-based Media: http://www.music.gla.ac.uk/HTMLFolder/Research/smatbamprivate/catego-ries.html, 1999-05-08, in [Boehm 1997 SmatBam].

<sup>&</sup>lt;sup>13</sup> Elementary or simple objects are objects made out of one entity or one binary (text files, bitmaps, wave format files, midi files). Composite objects consist of a number of elementary or composite objects, for instance a complex/composite music data structure. Complex objects are objects with attributes that change in size.

<sup>&</sup>lt;sup>14</sup> Examples for generic implementations and standard definitions can be found in OMG's Object Request Brokers and their Relationship Service Specification for distributed objects [OMG CORBA-Relation 1997] or in the Knowledge Interchange Format of the Laboratory for Advanced Information Technology [Finin and Labrou 1997].

- Inheritance one object inherits the characteristics of another object (e.g. all service provider users have read rights, these might be inherited down towards the developers of collections, who also have write rights; or, as a third example, all sounds stored at high quality inherit the characteristic of being served out over ATM network only).
- Association one object is associated with another object (e.g. Mendelssohn's composition Fingals Cave is associated with the geographical rock formation of Staffa. Another example would be that two pages can be associated with each other in form of a sequence. One page should follow the other in a certain context as for instance a book, course, slide show, score etc.).
- Attributes an object contains certain attributes, or certain characteristics which describe its content, state of being or its internal structure (e.g. all objects in the PADS archive have the attributes Title, Creator, Subject, Description, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, and Rights.<sup>15</sup>)
- *Web Links* Web-links can be thought of being a realisation of one of the above relationships in a web environment.

Thus digital resources, especially in the performing arts need a very powerful means of storing, managing and accessing intelligently the metadata to cope with the specialist problems time-based data contains. In addition to that, the metadata itself has to be powerful enough to support access to the structure of the data it describes. The Metadata has to be able to

- describe the whole resource (access, archive, re-use)
- describe the separate entities of a complex time-based resource
- have time-based or time-dependant description

The most efficient way to resolve this requirement is to manage the resources as objects in an objectoriented environment, which supports composite and complex objects and has in itself already inherently implementations of the relationships needed for any information management system. The metadata then can be used to "map" a knowledge domain, allow for flexible hierarchical metadata and metadata linked to points within time-based media data.



Fig.5 Examples of metadata use at the PADS. Metadata used for mapping a knowledge domain, mappings displayed graphically in three different views: a) as file manager, b) as graphical tree, c) as attributes an object.

#### 3.3. The Access

Having specified the users and the data, the one piece joining the two together is devising the access of the users to the data. The term "Searching" in this context has a meaning in its widest sense possible, implying the focused access or an active resource discovery. It should be stressed, that it is a matter of time, that common search and retrieval methods of finding books in on-line library catalogues

<sup>&</sup>lt;sup>15</sup> These are the basic Dublin Core attributes. One of the attractions of the Dublin Core metadata set is its simplicity - the Dublin Core was originally intended to be used by non-specialist authors to describe World Wide Web documents. The Dublin Core consists of 15 basic elements to which the AHDS workshop series and other initiatives from the library and information community have proposed some qualifiers and amendments to some of the definitions. [PADS Metadata 1997] [AHDS Metadata 1997]

or digital resources are enhanced by more intuitive searching/accessing technologies which take the above mentioned relationships, or rather their contexts, into account to form intelligent access technologies. This will be needed in a future, where the masses of information demand a very time-efficient way of focused access as well as a high success rate efficiency.

The matter is complicated by the users' contexts, such as specific specialist user communities with specialist needs. It can be stated that search methods will always depend on

- who is searching
- which resources is being searched for
- what archive is being searched in
- with which tools is the search supported.

Thus, besides traditional "search-and-retrieve approaches", there are others such as "metadata searches", "full-text searches", "closing-in-on-data-approach searches", the "browsing search", "graphical-, thesauri-, context – searches", "Demonstrate-and-Retrieve", "Content Searches", and of course any future, unimaginable searching technology.

Without going into detail of these approaches, which should spark the imagination by just their chosen names, it is clear, that in the best case searching should be based on the resource and its context (the object and its relationships), inclusively its use-context and not on the technology used.

#### 4. User, Data, Access and the MVC Paradigm

As detailed above, the three most important aspects to be taken into consideration for any digital libraries are very closely related to the specification and definition of the three-foldness of "users", "data", and "access". How the users is able to see, view and manipulate the data is controlled by the access, but is clearly very separated from it. In the computing world, this separation of content and representation has one of its object-oriented manifestations in the Model-View-Controller paradigm. The model being the content, the data, or a knowledge domain, the view being one possible presentation of it. The controller can be seen as the gadget maintaining the connection between the model and the view. One musical note, for instance, could be depicted in a system by an internal, proprietary data structure. To this note, one or more views can be "plugged in" as for instance a midi representation, a sound representation, and a graphic representation. Devising new views is thus independent of the content. See also [Ossenbruggen 1994].

This could be mapped seamlessly on the above three-foldness, the model being mapped without

any problems to the data, the view being the representation of the data, thus how the user will see the data, and the controller being how the model is interconnected with the view, i.e. the ac-

In a web environment, this concepts of model view separation has a vital importance, especially for services dealing with a multitude of data and

cess/manipulation/controlling means.

the management of them.

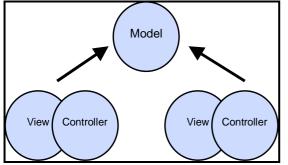


Fig. 6 Model – View – Controller concept, one model can have many controller – view pairs

Using this separation

- one object can have many views,
- one document can be represented in many languages,
- one website can be accessed through many user interfaces
- one collection can contain many access means
- one dataset can posses many representation
- one file can be delivered through different qualities of service.

Thus it may depend on the user and the collection, on how the resources are represented. A music user, searching just a collection of compositions, may see metadata such as "composer", "conductor", "first performance" displayed and with an high-performance connection is able to access sounds with CD-quality over high-performance networks, whereas on the other end a user from a general back-ground, searching through an interdisciplinary collection for any resource related to "Dickens", might see metadata displayed such as "creator", "role", "date of creation", "type of data", etc. Wanting to access some sound files, even though his connection might not be very fast , his delivery is achieved

via a lower quality of the material, thus keeping the same downloading/streaming speed, but with a lower quality of service. Below two data sets are shown, taken from two different collections on one database. Different metadata is presented for the two resources.

🚺 🚞 <u>Blackmail (Silent)</u> – 6 docs Type: film Format: mpeg Date: 1929: release Description: Story of a girl who kills a man in self-defence and is blackmailed by a man who know a play by Charles Bennett. Silent version. Credits: British International Pictures - production company, Alfred Hitchcock - director, John Ma; Frank Mills - assistant director, Alfred Hitchcock - adaptation Cast: Alice White = Anny Ondra, Anny Ondra's vocal double = Joan Barry, Mrs White = Sara Alig Charles Paton, Frank Webber = John Longden, Tracy = Donald Calthrop, The artist = Cyril Ritchar Hannah Jones, Chief Inspector = Harvey Braban, Detective Sergeant = Ex. Det. Sgt. Bishop, Gossi Monkman, The crook = Percy Parsons, The Sergeant = Johnny Butt, Harassed Underground travel Hitchcock Related Resources: Blackmail (sound) mpeg Related Resources: Blackmail dialogue script html 🚺 📾 <u>MacCunn: Land of the mountain and the flood, The</u> - 1 docs Composer: Hamish MacCunn

**Contributors:** orchestra: BBC Scottish Symphony Orchestra ; conductor: Sir Alexander Gibson **Type:** sound recording **Category:** orchestral

Date: 1989/07: recording, 1989/07: broadcast

Fig. 7 Two record sets in two different collections with different metadata displayed.

Within the database, the metadata might be stored in a certain format, such as "DC.creator = Alfred Hitchcock", "DC.creator.role = director" and "DC.creator = Hamish MacCunn", "DC.creator.role = composer". Displayed it shows a user friendly and discipline specific presentation of the metadata.

On another issue, with the separation of content from view, the implementation of different searching techniques for different collections, for different users or for different qualities of connections and client applications becomes possible. Thus a user may chose to browse through a collection via some subject index, or chose to graphically move through an VRML environment depicting his knowledge domain, or may chose to use one of the many search interfaces adaptable to his need of specialist resource discovery.

## 6. An example of a digital resource library

A goal of any digital library service is to provide interoperability with other collection holders by conforming to and implementing relevant standards. To shortly sketch the status-quo situation of using multimedia digital resource collections already available, one can look towards broadcasting stations, music/video archives, record companies and libraries. It must be taken into account that collections are stored in different storage mediums, ranging from simple file systems, to relational database management systems to the growing number of object-oriented database management systems. In addition, a large number of music catalogues in a variety of formats has to be also made accessible. Whilst in the academic and non-academic library world interoperability has established itself as an important topic, it is obvious that, commercial, television and broadcasting companies might not want their archives to interoperate with those of their competitors. However, given that material tends to decrease in commercial value with time but increase in academic and cultural or heritage value, it is quite possible that their material will end up in such a collection and so issues of interoperability are worth addressing, when planning any collection through its lifetime of existence. Besides already existing and longstanding library catalogue systems, with the prospective widespread use of digital resources for digital libraries, object-oriented database management systems will become a major means of storing, accessing and using complex, multimedia data objects.[Kahn and Wilensky, 1997], [Lagoze 1995] and [Lagoze 1996] Assuming a basic interoperability of different collections holding digital, multimedia objects, the underlying transfer protocol will have an influence on the performance, the quality and the representation means of the objects to be delivered.

It will be important to devise secure and distributed system, with collections stored in different locations, access handled from a central gateways and user access in the best case being controlled to a point of write, read and execute rights of single objects and collections. Solutions lie in the underlying existence of user rights management, such as a database management system able to control the access of many users in dependency of objects or collection of objects, or/and the use of a stateful protocols which are able to store session information as long as the users are accessing the resources. Below a host outline of PADS system architecture is shown, depicting very clearly the separation from data, delivery and view: The deepest level depicting different databases with their data, which is separated through a protocol conversion layer from the users, accessing the resources through different protocols. (see [Boehm, SMatBaM 1997] for detailed system specifications.)

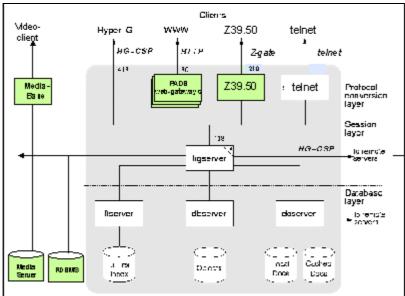


Fig. 8 System Architecture of the PADS

## 7. The future

There are a number of new technologies emerging, which will influence information access in a major way. MPEG7 emerged as a standard that might close one of the gaps between description and content. The Moving Picture Experts Group (MPEG ISO/IEC JTC1/SC29 WG11) is a working group of ISO/IEC in charge of the development of international standards for compression, decompression, processing, and coded representation of moving pictures, audio and their combination. In October 1996, MPEG (Moving Picture Experts Group) started a new work item to provide a solution for the urging problem of generally recognised descriptions for audio-visual content, which extend the limited capabilities of proprietary solutions in identifying content that exist today. The new member of the MPEG family is called "Multimedia Content Description Interface", or in short MPEG-7.<sup>16</sup> It will combine current technologies of automatic creation of metadata out of the content of a data with manually input data. Targeting business environments, sound signal processing technologies as well as image-signal processing and information retrieval technologies will be utilised to create a new standard for economically creating and using metadata without the need for lengthy archiving processes.

Other developments revisit specific metadata for specific metadata communities, such as metadata for educational material<sup>17</sup>, metadata for time-based media, such as the planned additions for SMPTE 12M-1995: for Television, Audio and Film - Time and Control Code, and as another similar development, creative user communities are starting to demand the existence of a creative metadata, the metadata which is being employed and created during the whole creation process. This metadata is currently discarded in most existent processes of putting a work into its final digital form.<sup>18</sup>

Specifically in music, new structured formats are evolving, some of which have inherently a metadata characteristic. Mpeg4<sup>19</sup> is an example of a very structured object-oriented multimedia format, SDIF<sup>20</sup> a emerging standard for describing structured sounds. Whereas in the music user community new structured music tagging languages are appearing, which could be called description languages, thus

<sup>&</sup>lt;sup>16</sup> see http://www.darmstadt.gmd.de/mobile/MPEG7, 1999-06-18.

<sup>&</sup>lt;sup>17</sup> http://www2.echo.lu/oii/en/meta.html#LOM, 1999-06-18.

 <sup>&</sup>lt;sup>18</sup> Carola Boehm et alii, CIRCUS Internal Rapporteur's Report of Group Alpha, Meeting Angouleme, France, April 99.

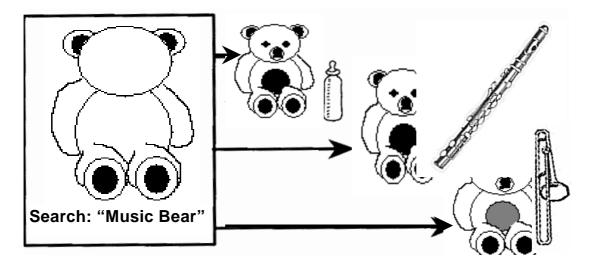
<sup>&</sup>lt;sup>19</sup> http://www2.echo.lu/oii/en/video.html#MPEG-4, 1999-06-18.

<sup>&</sup>lt;sup>20</sup> http://www.ircam.fr/produits/techno/multimedia/Cuidad/SDIF-e.html, 1999-06-18.

describing a resource well enough to represent the data itself. Examples for this are enhancements to  $SMDL/Hytime^{21}$  or  $GUIDO^{22}$ .

As a last notion of new metadata technologies, it can be expected that metadata will not stay textual data about data. Metadata in its meaning itself being "data about data" will be able to step into its multi-modal decade, in which a resource might be described by an abstract multi-modal representation of it. Thumbnails are good examples of imagery metadata of a resource, and similar technologies are already emerging in music formats, trying to filter out the most abstract and fast recognisable character-istic to browse sound collections. Theses might be images, or abstract filtered sounds, or a combination of both.

"Metadata is not something unique to the World Wide Web (WWW) and other forms of electronic data distribution. Metadata has been used for cataloguing and indexing information stored in libraries for over 2000 years. The introduction of computing has, however, vastly increased the speed and range of metadata searching."<sup>23</sup> Historically the data and the metadata was physically separate, one being stored in a catalogues, the other being stored in a shelf. But with above mentioned technologies, it is already possible that the separation of data from metadata or the description from the content ceases to exist. It will be then, that the weight of accessibility will not bear on the metadata, but rather on the structured data itself. Until then, we continue to describe our world verbally as best as we can, to define and to find the objects which we create for our working environments.



<sup>&</sup>lt;sup>21</sup> http://www2.echo.lu/oii/en/audio.html#SMDL, 1999-06-18.

<sup>&</sup>lt;sup>22</sup> http://www.informatik.th-darmstadt.de/AFS/CM/GUIDO/ 1999-06-18.

<sup>&</sup>lt;sup>23</sup> Martin Bryan on behalf of European Commission DGXIII/E, OII Guide to Metadata, http://www2.echo.lu/oii/en/metadata.html, 1999-06-18.

## References

[AHDS Metadata 1997] Greenstein, D. and Miller, P. (1997). *UKOLN/AHDS Metadata Workshop Series,* King's College, London (1997) http://ahds.ac.uk/public/metadata/discovery.html , 1998/07/01.

[Boehm, SMatBaM 1997] Boehm, C. & Malloch, S. (1997). SmaTBaM, Report on the evaluation process of system needs and demands of serving time-based media in the area of the Performing Arts, Glasgow 1997, http://www.music.gla.ac.uk/HTMLFolder/Research/SMaTBaM.html

[Finin and Labrou 1997] Finin, Labrou & Mayfield,(1997). *Laboratory for Advanced Information Technology: the Knowledge Interchange Format, KIF*, http://www.cs.umbc.edu/kse/kif/kif101.shtml.

[Flinn 1995] Flinn , Scott, (1995). Coordinating Heterogeneous Time-Based Media Between Independent Applications, ACM Multimedia 95 - Electronic Proceedings, San Francisco, California. http://www.cs.ubc.ca/spider/flinn/publications/mm95/scheduler.html.

[Kahn and Wilensky, 1997] Kahn & Wilensky (1997). A Framework for Distributed Digital Object Services, http://www.cnri.reston.va.us/home/cstr/arch/k-w.html, 1998/07/01,

[Lagoze 1995] Lagoze, D.& Shaw, E. & Davis, R. & Krafft, D. (1995) Erin Shaw,(1995) *Dienst/NCSTRL* http://cs-tr.cs.cornell.edu:80/Dienst/UI/1.0/Display/ncstrl.cornell/TR95-1514, 1998/07/01.

[Lagoze 1996] Lagoze, C. & Lynch, C. & Daniel, R. (1996), The Warwick Framework A Container Architecture for Aggregating Sets of Metadata.Cornell University.

http://cs-tr.cs.cornell.edu:80/Dienst/UI/1.0/Display/ncstrl.cornell/TR96-1593., 1998/07/01.

[Malloch and Pflicke, 1997] S. Malloch, S. Arnold, T. Pflicke, "Using Java to stream audio over ATM", Proceedings of the ICMC, Thessaloniki, 1997.

[OMG CORBA-Relation 1997] OMG (1997). *Relationship Service Specification for distributed objects*. OMG ftp://www.omg.org/pub/docs/formal/97-12-16.pdf

[Ossenbruggen 1994]. Ossenbruggen, Jacco van & Eliëns, Anton (1994). *Music in Time-Based Hypermedia*, Vrije Universiteit, Amsterdam. http://www.cs.vu.nl/~jrvosse/Papers/echt94/html/index.html, 1998/07/01.

[PADS Metadata 1997] Duffy,C.& Owen, C. UKOLN/AHDS Metadata Workshop : Moving Image Resources and Sound Resources, (1997) Warwick.

http://www.pads.ahds.ac.uk/padsMetadataWorkshopCollection

[Robertson 1997] Robertson, G. (1997). *Sample Rate Synchronization across ATM Network*, ICMC Proceedings of the International Computer Music Association, Thessaloniki.

[Robertson 1998] Robertson, G. (1998). *MIniMS, Multi-Participant Interactive Music Services*, University of Glasgow. http://www.music.gla.ac.uk/~george/minims/minis.html, 1999/05/08.