Proc. Int. Conf. on Dublin Core and Metadata for e-Communities 2002: 245-250

What's the Use of DC.Type? Semantic and functional aspects of the role of DC.Type within a moving image metadata generation tool

Simon Pockley
Australian Centre for the Moving Image
simonp@acmi.net.au

Abstract

Type has come to the fore as one of the primary organizing elements in the design of input forms suitable for the generation of high quality moving image metadata. A lack of semantic precision in both the definition and in the conceptual complexity of DC.Type's encoding scheme has prompted a re-evaluation of its usefulness as an element to be populated for interchange and discovery. In order to introduce precision to this element, a distinction is made between subject-based descriptors (genres), object based descriptors (forms), and manifestations or format-based descriptors (formats). A DCT2 vocabulary is proposed for DC.Type as a point of discussion for facilitating the deployment of domain specific encoding schemes and for filling gaps in the current list of terms.
Keywords: *DC.Type, type, genre, form, Semantic web, Metadata interoperability, Cultural heritage metadata, Knowledge management.*

DC.Type is one of the purest metadata elements, in so far as it is a term directly associated with our inherent need to order, categorize, classify and group similar resources. Yet, the benefits of semantic precision have been elusive for this element. Within the diverse reaches of the Dublin Core community, it has long struggled to find uncontested territory of its own. The many and various applications of the term, 'type,' have worked against it having a simple set of values. In its current form, sitting uncomfortably in the nether regions of the borders between DC.Format, DC.Relation and DC.Subject, 'type' could be an attribute of any element and therefore suffers from having to do too much.

DCMES 1.1 defines Resource Type as, the nature or genre of the content of the resource (being described). As it stands, the values of DCMI Type Vocabulary make up a coarse-grained but conceptually complex and semantically troubled list. This list compresses a single level of aggregation, several subjective descriptions or 'genres', and a mixture of high and low level physical format designators or 'forms'. The current approved list of terms consists of:

- Collection
- Dataset
- Event
- Image
- Interactive Resource
- Service
- Software
- Sound
- Text
- Physical Object (proposed)

The DC.Type Working Group's archives[1] trace the origin of this encoding scheme and provide fascinating insights into the various ways that reservation and important assumption can drop out when time is short and a schedule of deliverables takes precedence. There is no intention here, to be critical of the valuable work of this group. This group attempted to reconcile the natural, non-exclusive, non-hierarchical structures of usage with the forcing of unnatural, exclusive resource categories into a hierarchical classification scheme. The current encoding scheme is defended as a minimal, high-level list where low-level types can be extracted from domain, or application-specific, fine-grained terms. However, even at a high level, the useful application of such a complex mixture of terms is proving to be a semantically daunting task.[2]

In common usage, the term, 'type' is often used interchangeably with 'genre' and sometimes 'form' or 'nature'. It is used as a loose way of signaling a descriptive shift to a different level of aggregation. For example in a classical music web site that explains different musical forms, we can see a distinction between three levels of aggregation (my parentheses). The cantata is described as 'an important genre (level 1) of vocal chamber music':

> *Secular cantatas in German and Italian were composed by Keiser, Telemann, Bach and others, but this type (level 2) was never cultivated to the*

extent it was in Italy. In France and England the secular cantata was essentially an 18th-century genre, (level 3) emulating the Italian type (level 2). (Boynick 2001)[3]

In the arts, the various patterns of critical interests that have drawn on Aristotelian poetics as a way of aggregating works into types have achieved their status not because they fit together into any preconceived system or taxonomy, but simply because they recur constantly and independently. Literary theory is littered with the ruins of genre definitions that have convinced no one save their author. Communities of interest generally apply such terms as a pragmatic convenience where the act of categorization has occurred within a tradition of continuous redefinition.

The practical challenges of categorizing some of the more complex forms of new media that have appeared in the electronic, the digital, and the networked domains, are being addressed by communities that share an interest in managing the moving image in a range of analogue and digital formats. If, in the digital domain, they lag behind their text-based colleagues, it may be due to the twin challenges of complex technical dependencies along with massive file sizes.

Amongst members of the moving image community, as well as the movie consuming public, 'genre' has been, at once, the most useful method of grouping film and video, as well as the most deconstructed and conceptually unsound method of classification. This has arisen from an attempt to establish the credibility of media studies through an assertion of seriousness and separation from the less weighty entertainment values of Hollywood cinema. It has also come from a need to provide a grouping mechanism for the continuous production of top lists both as an aid to discovery and, by inference, to establish criteria for assessment or interpretation. After a century of film production, the conventions of 'genre' are also being used as stylistic shorthand as well as being an inherent component of the production of meaning. Like all literary forms, moving images constantly refer to themselves and to other cross-media generic manifestations.

In order to find firm ground on which to base a rationale for populating DC.Type, as an element with a consistent encoding scheme, it is useful to reach back into the origins of European thought and apply the triple distinction made by Aristotle between description by subjective response, by words, and by mimicry/imitation. This becomes a useful mechanism for distinguishing subject-based descriptors (genre) from object based descriptors (form), and manifestations or format-based descriptors (format).

This is not new. It is, in essence, the approach taken by Brian Taves (Chair) Judi Hoffman and Karen Lund in the *Library of Congress Moving Image Genre–Form Guide*. The guide uses MARC-based cataloging conventions to build up tri-part (genre-form-format) descriptions of moving image works. Notions of 'genre', and 'form' are described as follows:

Genres are recognizable primarily by content, and to a lesser degree by style. Genres contain conventions of narrational strategy and organizational structure, using similar themes, motifs, settings, situations, and characterizations ...

... Forms are defined as the basic categories indicating a moving image work's original exhibition and release parameters (such as length and medium), and which are separate from its actual content, not necessarily implying a particular narrative construction. Form terms include Feature, Short, Serial, Animation, and Television, and can be associated as needed with any genre, in a manner similar to free-floating subdivisions ... While the form indicates the work's original appearance, a third field, format, such as film, video, or videodisc, indicates the actual physical characteristic of any particular copy. For instance, a videodisc of THE SOUND OF MUSIC would have the genre-form-format heading "Musical—Feature—Videodisc". (Taves 1998)[4]

Responsive, non-linear forms might usefully be added to the Moving Image Genre–Form Guide. These would include such terms as: web-site, gameplay, generative, installation, interactive, simulation, surveillance, and ambient works. These are all forms that are (or can be) dynamic and open in nature. Library and archival communities have tended to avoid collecting examples of such works because they are difficult to capture except by 'snapshot'.

A useful test for 'form' is that form is an objective description with a precise but repeatable value. For example, a work described as a 'short' may also be an 'animation.' Whereas values for genre are imprecise, subjective terms with many shades of meaning that might be adapted to critical purpose such as documentary, film noir and crime.

Most of the values for the encoding scheme of DC.Type are, by this definition, high-level forms. Low-level precision will come with the ability to apply domain specific values for forms consistent with the notion of objective definition.

A semantic distinction between form and genre offers a level of precision that is missing in the approach taken in the *Metadata Object Description Schema* (MODS).[5] In this schema, 'type' functions as an element level attribute. For example, 'genreType' has form values: motion picture, newspaper, periodical, picture, video recording, web site etc.; 'typeOfResourceType' has form values: text, cartographic, notated music, sound recording, still image, moving image, three dimensional object, software, multimedia, mixed material etc.; and 'physicalDescriptionType has form as a subset along with 'internetMediaType' and 'extent' and is given the enumerated values of: Braille, electronic, microfiche,

microfilm (similar to DC.Format).

While genre terms might have limitations as subject heading values, communities who use and augment pragmatic applications of LCSH for discovery purposes would not find much difficulty in accommodating their own genre schemes into DC.Subject. Genre lists are by no means exclusive to moving images. Since 1991, Medical Subject Headings (MeSH)5 has listed 'Publication Types' to describe 'forms' of presentation. At its lower sub-type levels, genre terms are used to describe materials based on their cultural or literary forms such as addresses or sermons or their physical forms such as broadsides or posters.

At its higher 'form' level it is curious to note that the MeSH encoding scheme is considered to refine DC.Subject rather than DC.Type when they share terms.

Populating DCType

Type is a grouping attribute that could be applied to almost any DC element. When it comes to discovery, the challenge is to be able to extract information from rich records in a way that can be expressed using DC elements without disrupting inter-application interoperability. For practical discovery purposes (assuming an XML/RDF syntax), 'type', as a conceptual notion or display, rather than as a compounded element, could be retrieved or populated from a rich metadata record by an aggregation of the values of attributes from different elements. Assuming either DC.Type or DC.Subject had the benefit of the refinements of a separation between genre and form:

- DC.Type (domain vocabulary) Form + DC.Subject (domain vocabulary) Genre + DC. Format (domain vocabulary) Medium

or, through a qualified version of DC.Type:

- DC.Type (domain vocabulary) Form + DC.Type (domain vocabulary) Genre + DC. Format (domain vocabulary) Medium

For example, moving image, feature, DVD, or text, lecture, pdf

When the current DC.Type vocabulary was first proposed, the Working Group for DC.Type recognized that greater precision would be achieved by using more specific descriptors, but rejected the concept of multipart expressions on the grounds of 'retrieval considerations'.

> *... We expect additional structure for values of DC.Type to emerge from forthcoming discussion, allowing greater granularity of resource types to be expressed within this overall framework. This is likely mainly to involve sub-typing, for example including terms to indicate such things as moving vs. still images, different types of text, etc. However, the structure and syntax of Qualified DC has not been resolved at this time. A refined structure for Type will be implemented according to the general recommendations for Qualified DC. (Cox 1998)*[6]

Somehow between the separation of the minimalist approach expressed in Simple Dublin Core and the unrealized refinements of Qualified Dublin Core this form of semantic precision was lost to DC.Type. At DC9 in Tokyo, the DCMI Type Working Group decided that it would not try to produce an 'official DC' subtype list, and that such lists would be created by domain-specific working groups or by applications.

Conformance with the approved values of DC.Type involves transforming and extracting terms from lower level schemes and including the values of other elements.

What's wrong with DC.Type?

1. No encoding scheme registration process: A domain specific registration process is planned.

2. Image is too coarse: As a term, 'Image' needs some form of refinement. It currently covers any visual representation other than text such as photographs, paintings, prints, drawings, diagrams, maps, musical notation, animations and moving pictures, film, VR/3D environments and is a sacrifice to minimalism that compromises the usefulness of the term. The moving image is one of our major expressions of cultural heritage. At the very least still images and moving images should be separately defined.

3. Obscure terms: the term dataset (once data), as defined, stands out as belonging to the language of a technical community and might be replaced by 'template' as a term with more general currency.

4. Misplacement: 'Interactive Resources' may well have represented the zeitgeist of 1999 after the production of CD-ROM 'interactives' but in 2002 this term might be more usefully categorized as a subtype of a responsive or dynamic mode of encounter amongst terms such as:

- ambient works
- environments VR/3D
- game play
- generative
- installation
- interactive
- simulation
- surveillance
- web site

5. Element overlap: The term 'Collection' was proposed because of the need to identify a collection without describing its parts. Aggregations such as 'collection' are already expressed in DC.Relation.HasPart and one of the most important characteristics of

the Relation element is that it has an item level corollary Relation.IsPartOf. In practice, the term, 'Collection', by itself, is not nearly so useful because hierarchical trees or relationship models cannot be generated. Currently, the use of an implied default value to describe an item is obscure and an unnecessary complication to any encoding scheme. There is no question that the ability to describe a resource as a collection is needed. That this should be asked of DC.Type is worthy of challenge. The proposal to add 'Aggregation-level' to DC.Type by the DCMI-Government Working Group opens up the more complex issue of how to express levels of granularity.

From a moving image perspective, it is worthwhile noting that emerging standards such as MPEG 7 and MPEG 21 provide the syntax for describing sequence, shot, frame and even elements within the frame. With the aid of appropriate image recognition tools, these standards have the potential to turn all moving image items into collections.

From a discovery standpoint, the reason we aggregate works is to make it easier to get at the parts. In addition, usage of the term 'collection' is anchored in the Library or Museum community and confuses people who see themselves as building exhibitions, programs and packages rather than collections.

What's the use of type?

One of the unique exhibition spaces of the Australian Centre for the Moving Image (ACMI) is its Screen Gallery. This space, converted from two underground railway platforms into the world's largest digital media gallery, will feature the most innovative of Australian and international screen-based art, including:

- responsive installations
- large-scale projections
- video and computer animations
- interactive works
- net art
- immersive environments.

The juxtaposition or montage of film, television and multimedia will encourage multiple interpretations of themes, and an understanding and appreciation of how the various media interrelate.

The primary goal in outputting metadata conforming to standards such as the Dublin Core is to be able to exchange records with others and to expose selected fragments of this metadata for global exchange. These records also provide a source of content for footnote screens in the screen gallery, back-of-house administration, printed catalogues, reports, displays, lists, things to do, audio tours, interactive experiences as well as control over the flow of information about valuable assets (including the metadata itself). A significant departure from the item/format centred model of our main legacy database was to base the notion of what constituted a 'chunk' of information on the David Bearman model[7]. In this model, works are expressed in many forms and/or performed at many times and may be produced in numerous manifestations. Each metadata record is based on the intellectual content of the work rather than on its particular form, manifestation and format. For example a video postcard work by Robert Cahen entitled Cartes Postales can be expressed in a linear form as a short and be manifested as VHS video in PAL or it could find expression as a non-linear multi-screen two-channel installation in MPEG2 at 6 mbs.

The changes that networked digital technologies have made in the way digital content can be produced and, by implication, discovered and consumed are most evident at the point of creation. They have already resulted in some significant changes in the management of audio-visual content, regardless of format.

1. Shift from passive consumption to active use/production

Cheaper digital moving image production tools (such as the iMac) combined with in-built encoding software are leading to increased screen literacy with an explosion of rich media content. We are also beginning to see tools with meta-logging software built in to the production and editing cycles (e.g. Sony Tele-File). It is important that such tools and content management systems are flexible enough to be able to be integrated with other systems. Yet the end-to-end approach of vendors forces a proprietary dependency anathema to collaborative or distributed activities and metadata exchange.

2. Shift in managing multiple manifestations

The re-purposing of rich media content goes beyond proprietary obstructions to cutting and pasting combinations of audio-visual fragments. Often multiple manifestations are required of a single work to suit different outputs and configurations. In the face of rapid developments in encoding software, it is important to attempt to store master files of uncompressed content from which different encodings can be made. Such content is unlikely to be exposed for public consumption. Depending on your point of view, rich media outputs are often manifested in unsuitable formats such as film, video, low-resolution codec, proprietary and even redundant formats.

While many cultural institutions are embarking on expensive digitisation projects for legacy content it would make sense to know who else holds the same resources and if they have already been digitised.

3. Shift in identifying the borders of the work

In a primitive way, the web has created a contextual universe around almost anything we can identify

with text. We now expect to investigate resources related to works that we may have previously viewed in isolation. The placing of borders around chunks of content has become a source of contention, about to be compounded by the wider deployment of RDF. The borders of a work have become as conceptual as the notions of what constitutes a collection.

Similarly, the reach of an Application Profile may soon define the borders of a business or a cultural or educational institution. In such spaces contextual resources are as inseparable from the notion of the work as the idea of it - in space and time.

These changes combine to create combinations of rich and domain specific metadata schema suitable for discovering complex digital resources. We presented a paper at DC9 in Tokyo outlining some of our experiences and practical difficulties encountered in the collaborative cataloguing of a wide range of digital artworks. Since then, the 'buy in' of curators and programmers has come through the development of different 'views' of our metadata generation engine or catalogue. Members of the D.C. Community who have had experience within cultural institutions will understand that exhibition oriented Curators and Programmers (key metadata creators) have quite different views of resources than Collection Registrars, Librarians and Conservators.

Our main cataloguing tool is a metadata engine that adapts itself to the perceptions and language of a range of users by providing them with different views of the record and its component outputs.

In an ideal world, the generation of high quality metadata begins at the point of creation. However, ACMI is a cultural institution that engages in creating or producing exhibitions and programs, commissioning works; and acquiring works by donation, purchase and internal production. This means that the process of metadata generation begins at the point of accession. For donated collections and failed production encodes, this can sometimes mean that the first metadata created is actually a record of deaccession. The point of accession or ingestion or disposal can vary according to whether a work is entering or leaving our collections.

A view of the record, tailored to the inputs needed to complete it, is activated by the selection of an Accession type from an administrative schema.

- Exhibition (a collection created by Curators)
- Program (created by Programmers either collection or item level)
- Event (created by Programmers either collection or item level)
- Production (internal – either collection or item level)
- Commission (external – either collection or item level)
- Purchase (includes donation either collection or item level)
- Loan (either collection or item level)
- Disposal (either collection or item level)

ACMI's Application Profile uses an XML/RDF syntax to augment and populate Dublin Core elements from a range of fine grained elements and attributes relating to the cross referencing of:

- Descriptive metadata: textual and visual documentation e.g. clips, stills, artist's statements etc.
- Interpretive metadata: e.g. exhibitions, programs, rationales, curatorial statements, interpretive essays, reviews, genres etc.
- Expressive metadata: technical requirements e.g. equipment lists, instructions, layout plans etc.
- Physical/production metadata: format and display descriptors e.g. aspect ratio, resolution, signal, frame rate, audio encoding, bit rate etc.

Terms selected from an unapproved DC.Type encoding scheme are used as triggers for displaying the appropriate elements to be populated. In our case, the conditional use of the high level form, 'moving image' can determine the values of attributes needed for recording complex Format descriptors for over 90,000 titles.

The categories of information needed to manage a range of MPEG2 moving image manifestations are quite different from those required for the time and place of an event which has no format; or text; or the dimensions or location of a physical object.

Reworking DC.Type into DCT2

This paper acknowledges that different domains have quite different terms and needs. However, as a way of addressing some of the more restrictive consequences of an hierarchical approach to arranging the values of an encoding scheme, usage and broad representation would suggest several small but pragmatic changes to DC.Type. The following alterations would greatly improve the useful application of DC.Type and the consistency of lower-level encoding schemes:

1. an adjustment to the DC.Type definition to replace the word 'genre' with 'form' where form is described as an objective description of the resource;
2. an adjustment to the DCT1 encoding scheme to include only terms that are forms;
3. splitting the term image into the two high level terms, 'still image' and 'moving image'. This may require the DC.Type encoding scheme DCT1 evolve to DCT2 where the moving image and the still image are recognised as distinct top level terms with the definitions:
   moving image: *Definition: Any image created in a film, video, or other media format that alters with time and that is able to be displayed or projected on a screen. For example, movies, animations, television, multimedia, games, emerging media, simulations.*

Table 1. Table of proposed changes

Current DC.Type scheme	Proposed DC.Type scheme	Rationale
Collection		Resolve to DC.Relation.HasPart
Dataset	Template	More common usage suggestion
Event	Event	
Image	Still Image	
	Moving Image	
Interactive R.	Responsive Resource	Suggested (possibly dynamic)
Service	Service	
Software	Software	
Sound	Sound	
Text	Text	
Notation	Suggested possibility	
Physical Object	Physical Object	
	Web site	Added term

still image: *Definition: the content is primarily symbolic visual representation other than text. For example - images and photographs of physical objects, paintings, prints, drawings, other images and graphics, diagrams, maps. Note that image may include both electronic and physical representations.*

1. adding the term 'web-site' as a top level term;
2. removing the aggregating term, 'collection' from the scheme and resolving it within DC.Relation as a term that need not have its parts described;
3. recognizing that 'interactive' is now a lower level term of a form that is responsive or dynamic;
4. expediting the registration process for domain specific encoding schemes.

[1] DC.Type Working Group archives 1999-2000. http://www.mailbase.ac.uk/lists/dc-type/archive.html

[2] Agnew, Grace. V-Access@Listserve.UTK.EDU message sent 2002-05-31 thread: difference between 'genre' and 'type'? *... Somehow or other we never ended up discussing this recommendation fully in our user guide. We also didn't see a useful place to put genre because, on the one hand, while it can be subject, our scenarios of use genres really aren't subjects. DCMI seems to use genre as synonymous with format, and that wasn't appropriate. Finally, we settled by default on the catch-all data element, "description".*

[3] Boynick, Matt. The Classical Music Pages: Musical Forms – Cantata. Last Revision - 10 October 2001 http://w3.rz-berlin.mpg.de/cmp/g_cantata.html

[4] Taves, Brian et al. 1998 The Moving Image Genre-Form Guide. Library of Congress Motion Picture/Broadcasting/ Recorded Sound Division, February http://www.loc.gov/rr/ mopic/migintro.html

[5] Metadata Object Description Schema (MODS) http://www.loc.gov/standards/mods/

[6] Cox, Simon et al. Type Element Working Draft. 1998 http://www.dublincore.org/documents/1998/10/23/type-element/

[7] Bearman, David et al. 1999 A Common Model to Support Interoperable Metadata. http://www.dlib.org/dlib/january99/bearman/01bearman.html