AEDM

ARTICLE

Developing a Metadata Standard for Multimedia Content Management: A Case Study

Abdullahi Mohammed Sheriff[1]*, Dino Bouchlaghem[1], Ashraf El-Hamalawi[1] and Steven Yeomans[2]

[1]Department of Civil and Building Engineering, Loughborough University, Leicestershire, UK
[2]Buro Happold Consulting Engineers, Bath, UK

Abstract

Metadata is considered crucial for the organization, management and retrieval of data within a content management system. While numerous works exist that define the critical nature and importance of metadata, little practical guidance is available with which organizations looking to develop metadata standards to meet their corporate needs can do so. This article presents a case study in which a metadata standard was developed for an international construction industry-based consultancy for managing content within a proposed multimedia library. Three parallel activities were carried out as part of the research methodology. A desk study was initially conducted to capture the metadata used in the previous library. A review of 11 metadata standards was carried out to obtain a baseline of suitable attributes. A workshop was then conducted with a sample of end users to further capture specific requirements. The outcomes from all three exercises were then analysed to obtain a company-wide metadata standard. A further refinement was carried out to rationalize the list into a core set of attributes. A closeout workshop was then conducted with key participants to identify lessons learnt and review the outcome of the project. Drawing from these, this article further adds to knowledge by proposing a 12-step guide to enable organizations develop similar metadata standards to meet their needs. The research outcome also shows that while existing metadata standards can be used as a starting point, no specific standard is comprehensive enough to meet the needs of an organization without appropriate levels of customization.

Keywords – content management; enterprise content management; information management; metadata; taxonomy

INTRODUCTION

UNDERSTANDING METADATA

Metadata refers to additional pieces of data or attributes that describe the context and structure of a piece of data, content, document or other bits of information and their management through time (Bjork, 2001; ISO 15489: 2001). The use of metadata within a content/document library allows information to be more easily found, its source determined and its context understood easing interpretation and enabling re-use (Bentley, 2001; Rockley et al., 2003; NISO, 2004; Day, 2006). While resource discovery remains one of the principal functions of metadata, others may include provenance, technical specification, functionality, administration, content ratings and demonstrating linkage or relationships (Technical Advisory Service for Images (TASI), 2006). In order to fulfil these functions, metadata should be seen to be application-independent, clearly defined enterprise-wide and used as a tool to facilitate

*Corresponding author: *E-mail*: sheriffson@hotmail.com

doi:10.1080/17452007.2011.594569 © 2011 Taylor & Francis ISSN: 1745-2007 (print), 1752-7589 (online) www.earthscan.co.uk/journals/aedm

interoperability between multiple systems as organizational needs warrant (Bjork, 2001; Haynes, 2004). Metadata can generally be seen to be of three types (NISO, 2004; TASI, 2006):

1 Descriptive metadata, which are used to find, identify and understand a piece of content examples including the title, abstract, author and keywords.
2 Structural metadata that show and define relationships and associations between content or compound content made up of smaller bits of content, an example of which is relation.
3 Administrative metadata aimed at managing the lifecycle of content and associated technical information, examples of which include date created, file type, file size and restrictions.

As Burnett et al. (1999) explains, in the end, metadata should enable users to clearly answer two principal concerns, that is what information is available and what information is useful.

EXISTING METHODOLOGIES FOR DEVELOPING METADATA STANDARDS

Content management systems, whose key feature is descriptive metadata, are increasingly being used for managing electronic content within organizations (White, 2005). While international standards for descriptive metadata attributes do exist, research aimed at defining metadata standards are predominantly bibliographic, with very limited research on organizational metadata (Karjalainen et al., 2000; Murphy, 2001; Paivarinta et al., 2002). In particular, no clear methodology exists with which organizations can develop appropriate standards to meet their needs. One such exception is the detailed case study presented by Paivarinta et al. (2002), in which a metadata standard was developed for an enterprise wide Electronic Document Management System in Fortum, an international energy company. Similarities do exist between both studies, particularly in the methodologies and the metadata attributes identified, all of which will be highlighted below.

Despite the shortage of research in this area, some process methodologies for developing broader information management strategies (which may include metadata) have been proposed, the most prominent of which is the designing and implementing record keeping systems (DIRKS) framework. The DIRKS methodology is a structured eight-step process that provides guidance to organizations on how to improve information and records management (National Archives of Australia (NAA), 2001). The eight processes defined can be understood thus: Step A contextualizes the project and focuses on understanding the company, its business structure, nature of the market and industry regulation. Steps B and C gather specific record-keeping requirements and match with the business activities of the company. Step D then focuses on a state-of-the-art review of the company, assessing the appropriateness of current record management policies, systems and procedures to meet the requirements gathered. Where deemed inappropriate, this is then followed by Stages E and F, aimed at redesigning the entire strategy to address identified limitations and meet the company's strategy. Implementation and post-implementation reviews are then carried out in Stages G and H to complete the model's lifecycle (NAA, 2001; Hofman, 2006).

The DIRKS model is holistic, incorporating aspects of electronic systems, human interaction and organizational behaviour (NAA, 2001). Although developed for the purposes of record's management, the steps outlined can be adopted for developing broader organizational information management strategies. The gap in the DIRKS model is that while metadata are acknowledged to be crucial in supporting the functionality of the records management system, it explicitly provides no specific guidance on how organizations looking to develop a metadata standard either for a specific project within an organization or across the whole organization can do so.

Bock (2005) also proposes a simple three-step process for designing metadata with particular emphasis on metadata for digital asset management systems. The three steps begin with *capture*, which involves aggregating and understanding the unique terms used by the target groups to describe digital assets in their respective business processes. Following this is *curate*, which involves a refinement of the set terms including the similarities and

differences among them and determining which ones are important and why. The final step is *categorize*, which involves defining the metadata sets that will be used for each group. This methodology, while simple, is focused entirely on the actual identification of metadata for a project and does not take into account other critical aspects such as the implementation, or even adoption, of descriptive metadata standard attributes. The methodology was therefore deemed inadequate and not holistic enough for developing a company-wide metadata standard and a long-term framework for its subsequent maintenance. Thus, it could be said that despite these and similar frameworks, there remain limited references for practical guidance as to how organizations can develop metadata standards to meet their needs.

This research was carried out as part of a four-year Engineering Doctorate programme (EngD) into Enterprise Content Management (ECM), with a focus on developing an information management strategy for construction-industry-based organizations. This article presents the outcomes of an exercise to define a metadata standard for a proposed multimedia library within an organization (as a detailed case study). It begins by introducing the aim of the research and the organizational context. The methodologies adopted are then presented, accompanied by the details of activities carried out. This is followed by a discussion on the findings and outcomes of the research, along with the lessons learnt. Based on these, a step-by-step guide is proposed through which other organizations can similarly define metadata standards to meet their needs. It is anticipated that this guide, although focused on metadata, can and should be used to complement broader and more comprehensive information and records management frameworks such as the DIRKS model. For reasons of anonymity, the case company involved in this research will be referred to as Company A.

THE PROJECT SCOPE AND RESEARCH CONTEXT

Company A is an international construction industry-based engineering design consultancy employing over 1,800 staff in over 25 offices across the world. Serving multiple geographically dispersed projects, the company relies on information-sharing and collaborative working to ensure effective project delivery to meet clients' needs. To support core business processes (including marketing and bid management), a new multimedia library was required. Prior to this, a bespoke image library existed, designed by an 'in-house' IT team approximately eight years ago. Being bespoke, it was designed with suitable functionality to meet the user's needs at the time. While the functionality was still considered relevant, the system had since developed some significant shortcomings.

As the company evolved both technologically and as a business, the use of higher resolution image formats and other multimedia file types such as videos, etc. have become prevalent, all of which the existing system with its functional, technical and storage limitations could no longer manage. To overcome this, users began to manage multimedia content in personal hard drives and project servers. This brought with it business risks that Company A was keen to mitigate. Using such compartmentalized storage solutions also meant that content stored in one environment was inaccessible to other staff that may have required it to fulfil their tasks (thus resulting in more business inefficiencies). The existing library also had limited search-ability and indexing capability. Its' simple functionality asked for users to provide very limited amounts of metadata when uploading content. Although this was complemented by the use of keywords added by users, retrieval of the required content was difficult, which in turn negatively impacted the usability of the library. The less the users were able to find content, the less confidence they had in the system and therefore lesser the likelihood that they would use the library in future to either store or retrieve images.

In replacing the library, Company A also wanted to improve its overall workflow capability to optimize business processes. Rather than simply store content, the requirement was to provide a capability for assembling information products (i.e. documents, reports, bids, etc.) 'on-demand'. Currently, the process for carrying out such tasks requires an Adobe in-design document to be created, with the

images saved to a local folder from which they are embedded into the document. The document is then sent to relevant parties in a zip folder for approval (with copies created). The favoured content-on-demand approach would aim to reduce associated storage costs, optimize the flow of information, enable standardization and save time. It was realized that core to meeting these requirements is the development of a robust metadata standard, particularly to support search and retrieval, the principal thrust of this project.

RESEARCH METHODOLOGY

The objectives of this research were: to define a metadata standard for information management and retrieval within a case study organization; investigate issues that emerge in the metadata development process; and identify a generic methodology that may be used to guide practitioners in creating appropriate metadata standards. A case-study approach was adopted due to the depth of understanding sought with regard to specific company business processes, the holistic investigation of the concepts required and the company-specific scope of the project. Fundamentally though, this research is exploratory in nature, seeking to explore patterns inherent in a specific activity carried out in a real-life context (Yin, 2002). Thus, although a framework is proposed, no attempt is made at generalizing the findings, as it is acknowledged that the sample size is not sufficient to do so. The project was conducted by the researcher alongside a team of user representatives and company stakeholders. These were sourced primarily from those disciplines across the company deemed to be either the primary users of the system, key secondary users or those actively managing the system. A total of eight such disciplines was identified with two others invited to join, on the basis of their additional expertise that was considered to be vital to the project. Care was taken to ensure broad user-requirements across the company were sufficiently represented. At some stages, input from other users (outside of the group) was sought to either re-confirm the findings or gauge opinions on specific issues. Three complementary approaches were adopted in line with the objectives of this project. These were: a desk study of the existing library, a literature review of existing metadata standards and a workshop with user representatives.

The desk review of the existing library was carried out to analyse the existing structure and metadata (if any) being used. This review was essential for obtaining a better understanding of the nature of the problem and potentially to identify any embedded patterns in the way images were currently tagged. The findings were then mapped out and analysed using a mind mapping tool. Since the existing library did not require strict adherence to any attributes, it was presumed that users would tag files in manners that made the most sense to them or would enable them to retrieve it best. This exercise was focused on identifying user metadata preferences; thus the frequency with which certain attributes were used was deemed unimportant and not captured. To complement the desk review, a parallel review of existing metadata standards was carried out to determine the suitability of existing standards to meet the needs of the company. This review was carried out in close consultation with the company stakeholders and user representatives. Based on user requirements, the scope of the project, the peculiarity of business processes and terminologies within the construction industry, no standard was identified to be wholly suitable to the needs of the company. An analysis was then conducted to identify attributes that were common to their standards, to establish a baseline from which an appropriate standard could be developed. Third, a workshop was held with the project team and selected user representatives, to identify their specific metadata needs and also to compare the findings with the outcome of the two exercises performed earlier. During the workshop, 14 images and videos were selected at random from a pool of 4,000 and presented to the user representatives. As for each image, representatives were asked what metadata they would like to attach to the files (no suggestions were offered). The outcome of the previous two exercises was not communicated to them until the end of the workshop. After the workshop, all metadata suggestions were aggregated into a spreadsheet. As with the desk

study, the object of this exercise was not to identify the frequency with which certain attributes occurred but to capture all the attributes required.

A conscious attempt was made to ensure that the metadata standard developed was driven by user needs. Thus it was felt that no individual exercise from the above could be carried out in isolation and deemed to comprehensively reflect those needs. Due to the limitations of the existing image library and the fact it was no longer used by many groups across the company, the results of the desk study would be fairly limited. Studying the standards alone would also be inadequate, as it did not and could not reflect the organization's specific requirements and business processes. The workshop, when accommodating user preferences, would not have helped facilitate interoperability between systems, as it would lead to exclusively bespoke solutions to meet only the needs of this project. Thus all three exercises were necessary, providing the breadth needed to address the limitations of each singular exercise. To refine the findings, a comparison was then conducted on the collective outcomes of the exercises. Attributes featuring in two or more of the three exercises were considered to be important and retained, whereas attributes featured in only one required justification by the end users as to its perceived importance. Where such justification was not provided, the attribute was removed.

Having obtained the standard, it was then observed that the full deployment of all the attributes may create a system too burdensome for the end users, thus impacting its effectiveness. A further refinement was therefore carried out by the project team, by re-organizing the attributes into three tiers based on their perceived importance (tier one attributes were deemed critical and tier three attributes were to be deemed 'nice to haves'). Prior to implementation, a further analysis was also carried out to identify which of the attributes could be automated and which would have to be manually input by the end users, helping in further reducing the total number of attributes and completing the project. A project closeout workshop was then held, with the project team to critically review the outcome of the project, the steps undergone and the refinement process. Details from each exercise and the outcome of the research are outlined below.

THE PROJECT

DESK REVIEW

The structure of the current image library with its nearly 4,000 images was studied in detail to reveal the following 15 attributes:

1 *Date* – The date at which the activity/event was carried out. In a project, the date reflects the date of specific activities ranging from early-stage preliminary site visits to project-closeout visits and even post-occupancy appraisals. It provides an audit trail and useful context for the content of the image.
2 *Contributor* – The person/entity responsible for uploading the image into the library. This sometimes differs from the person that took the actual image (see below) and is always a company employee.
3 *Source* – The person/entity credited for supplying the image. This could be either a member of the company's staff or an external photographer.
4 *Creator (or Photographer)* – The person/entity that created the image, sketch or drawing. This may sometimes be the same as the source, but where sources were secondary, the creator often differed.
5 *Rank/position* – This was used in images relating to people and/or events. It relates to the designation of the individual(s) that appears within the image. These were either single individuals, in which case specific titles were used (e.g. HR manager), or a group of individuals, in which case a more collective title was used (e.g. graduate engineers).
6 *Group/team* – The discipline within the company to which the content of the image was attributed. For example, certain project images with pile foundations were also tagged with ground engineering (the group name responsible for or affiliated with the image).
7 *Office* – The specific office of Company A associated with the content or subject matter of the image (and not the image itself).

8 *Business region* – The specific business region in which the office (as described above) is based and/or the business region with which the content of the image is associated.
9 *Keywords* (e.g. winter snow, Plug socket, etc.) – Keywords were the most used attributes in the existing library, as they allowed users to specify contextual terminology related to the content, which would enable them retrieve it. The keywords used varied from subject-technical scientific terms including materials, etc. to general descriptive terms such as forest, summer, sunrise and so on.
10 *Location* (including city, country and continent) – The geographical setting of the content within the image or the activity/individual associated with the image. In some instances only the city's name was used whereas in others that of the country and, indeed, of the continent was used. All geographical references in this instance were prepared using GIS co-ordinates.
11 *Designer* – The individual/company or entity directly credited with creating the content captured in the image. This was predominantly used for buildings, infrastructure, project images, monuments and similar physical structures.
12 *Content type* – These are categories that reflect the genre of the content. Examples of these include sketches, drawings, presentations, maps and so on.
13 *Format* – This relates to the format of the image. This was automatically captured by the image library, and it is unclear if users relied on it for image retrieval.
14 *Resolution* – The resolution of the image was used as a core metadata attribute with which high-quality images were identified and retrieved.
15 *Access rights* – This relates to the sensitivity of the image and the permissions that indicate those that should have access to it.
16 *Copyright* – This was included as part of company policy to ensure that regulatory requirements are met and that intellectual property rights were never infringed upon.

The existing image library was restricted in the number and nature of attributes it allowed users to add; thus, the findings from this desk study were incapable of being wholly reflective of the users' needs. Also, within both the project and generic sections, the library used a system of categories and sub-categories to form two levels of hierarchy. No specific definitions were provided, either to the researcher or to users uploading content into the library, as to what the terms 'category' or 'sub-category' meant. This exercise observed that there was no apparent consistency in their use. What was apparent was that attributes used in the hierarchy reflect similar attributes employed as metadata. Indeed, some metadata attributes, such as date of the event/activity, were used frequently among the sub-categories. As an example, consider sub-category 'RedR day 9/12/05, 2005' (under the category 'company events', RedR day is a company event). This reflects both the name of the event/activity and the date. In all such cases both attributes were counted as a single instance of metadata.

REVIEW OF EXISTING METADATA STANDARDS

IDENTIFYING THE STANDARDS

A review of metadata standards was then carried out to identify if any directly met the needs of the company and whether they could be wholly adopted. Where such standards could not be identified, the exercise then aimed to develop a baseline set of metadata upon which the company standard could be built. As part of the standardized information architecture the company favours, preference was given to the use of external metadata standards instead of developing bespoke solutions. Preliminary research suggested that no standard can be applied without modification to suit a particular company's needs (Paivarinta et al., 2002; TASI, 2006). Two previous studies were carried out on the basis of a similar review of available metadata standards. Paivarinta et al. (2002) conducted a similar comparison, using 18 metadata standards, whereas an earlier study by Burnett et al. (1999) was based on a comparison of six standards. As it was intended that this exercise would build on the baseline defined by Paivarinta et al. (2002), care was taken to identify standards that had not been used in their research. This was to ensure that no standards

were repeated (thus duplicating the results) and that new perspectives that have emerged, since both studies were carried out, are accommodated. Only the ISO15836 included in both studies was also included here. This is because it has emerged as the de facto standard for descriptive metadata. The following standards were reviewed:

- *ISO 15836* – This is based on the Dublin Core, a ubiquitous standard for cross domain resource description, often referred to as the de facto standard for descriptive metadata (ISO15836: 2009).
- *Visual resources association (VRA) core* – A metadata standard for describing images and visual content for the cultural heritage community (TASI, 2006).
- *E-government metadata standards* – This lists the elements and refinements used by the public sector in the UK to create metadata for information resources. It also gives guidance on the purpose and use of each element (e-Government Metadata Standard, 2006).
- *E-records* – A metadata standard developed specifically for the purpose of effective records management in the public sector in the UK, it is built on the e-government metadata standards (The National Archives, 2002).
- *ISO 19115* – An international metadata standard for describing geo-spatial datasets (ISO 19115: 2003).
- *BS-1192* – Set of standard procedures and methodologies for managing the production, distribution and quality of construction information, using defined processes for collaboration and specified naming procedures (BS 1192: 2007).
- *UK learning object metadata (UKLOM)* – A standard for the interoperable description of learning objects that is, i.e. any entity (digital or non-digital) that may be used for learning, education or training (IEEE, 2002).
- *New Zealand Government Locator Service (NZ-GLS)* – A standard metadata element set designed to improve the discovery, visibility, accessibility and interoperability of online information and services in a cross-disciplinary information environment in New Zealand (Archives New Zealand, 2004).
- *UK-Gemini* – A defined element set for describing geo-spatial discovery level metadata within the UK (Cabinet Office, 2004).
- *Australian Government Locator Service:* A standard metadata element set designed to improve the interoperability and retrieval of online information and services with primarily in Australia (National Archives of Australia (NAA), 2006).

Taking into account the 18 standards already reviewed by Paivarinta et al. (2002), it could be said that the baseline developed for this research reflects a secondary review of up to 27 standards.

THE ATTRIBUTES

Attributes that occurred consistently in over four of the now-11 standards (including the review of Paivarinta et al., 2002) were identified and earmarked as the baseline standard. These were given in Table 1.

This baseline is similar to that established by Paivarinta et al. (2002), as all but two of the attributes specified in their work were still dominant here. The absent attributes were keywords and

TABLE 1 Baseline developed from review of existing standards

NO	ATTRIBUTE	NUMBER OF INSTANCES
1	Creator/originator	11
2	Title	10
3	Description/notes/abstract	10
4	Date	10
5	Type	10
6	Format	9
7	Relation/lineage	9
8	Accessibility/availability/rights	9
9	Subject	8
10	Identifier/ID/drawing number	8
11	Language	8
12	Source/supplier	6
13	Coverage	6
14	Publisher	5
15	Contributor	5
16	Location	5
17	Status	5

organization. Three additional attributes not reflected in their work also emerged. These were: source/supplier, coverage and status. A comparison with ISO 15836 (2009), the most widely used descriptive metadata standard, showed that all but one of the attributes in the standard are reflected here, with only two attributes, that is location and status, absent from it. The resulting baseline reconfirms the statement, of Paivarinta et al. (2002), CEN (2005) and TASI (2006), that the Dublin core metadata can be suitably used as an established baseline from which organizations can build their own metadata standards. However, a review of this baseline by the company project team clearly showed that there were other attributes specific to internal business processes and to the construction industry that were not reflected here, but which the users considered to be critical in effectively carrying out their tasks. Hence this did not reflect a comprehensive list of descriptive metadata suitable to the needs of the company.

THE WORKSHOP

To capture end-user-specific preferences, a workshop was set up in which user representatives from the different disciplines were provided with 14 randomly selected images and videos sourced from the existing library that yet reflected the breadth of content. The attendees for the workshop were the user representatives described in the 'Research Methodology' section . In selecting the images and videos to use, the only criterion employed was to ensure that they were broadly reflective of the variety of images within the library. The participants were all in the same room and were allowed to openly discuss their thoughts with each other throughout the process. As the aim of this exercise was to aggregate all the desired attributes, no attempt was made to tally the number of times a single attribute was highlighted or to evaluate its level of importance. The results from the previous two studies were not shared with the user representatives prior to this exercise, to avoid subtle influences on their choices. In the end, a total of 56 attributes was identified. A subsequent refinement was carried out to remove duplications, synonyms and attributes that were considered to be subsets of other attributes. Justification was also sought for each attribute, the results of which are presented below.

FINDINGS

COMPARISONS OF THE THREE STUDIES

In analysing the collective outcomes, a comparative study was carried out between the three activities. It was observed that although the outcome of both the desk study and that of the baseline developed were not revealed to the user representatives prior to the workshop, considerable similarities emerged between the outcomes, as reflected in Table 2. Certain attributes matched each other in definition but differed in terminology. In these instances, the term deemed more easily understandable by potential users was chosen. In two instances, the term differed from those adopted in the standards. Differing from the recommendations of TASI (2006), it was believed that using terminologies adopted in standards but unknown to the user would render the problem more difficult for users. One such term was 'coverage', defined as the extent or scope of the content of the resource (TASI, 2006). This was similar to file size. Instead of either definition, 'resolution' was adopted, as users showed preference for images, not on the basis of its size but on that of its clarity and level of detail. 'Relation', which featured in the standards reviewed, also matched 'related press articles'. However in this case, the former was used as its meaning was flexible enough to include a relation to other media or content outside the press articles suggested in the latter's definition.

For each attribute reflected in only one of the three exercises, a case-by-case justification was sought for its inclusion. Numerous such cases emerged from the workshop, as it was the principal medium through which company-specific requirements were identified. Where the justification was deemed adequate by consensus, the element was retained and where it proved inadequate, it was removed. 'Description' occurred only in the standards reviewed (one of three), but it was included as it was deemed important. The language and the publisher were excluded as they were deemed inapplicable. 'Subject' was deemed to be reflected in the

TABLE 2 Comparison of metadata attributes

	ATTRIBUTE	WORKSHOP	STANDARDS	DESK REVIEW	
1	Access rights	X	X	X	
2	Bid outcome	X			Considered important therefore included
3	Business region	X		X	
4	Company	X			Considered important therefore included
5	Content type	X	X	X	
6	Contributor (of image)	X	X	X	
7	Copyright	X		X	
8	Creator	X	X	X	
9	Date of event/activity/project	X		X	
10	Date image taken	X	X		
11	Date of commissioning (or use)	X			Considered important therefore included
12	Drawing no	X			Considered important therefore included
13	Designer (architect)	X		X	
14	Description		X		Considered important therefore included
15	Duration for video/coverage for images	X	X		Considered important therefore included
16	File format	X	X	X	
17	Groups/disciplines	X		X	
18	Identifier		X		Not included – auto software generated not descriptive
19	Interesting space	X			Considered important therefore included
20	Key company staff	X			Not included
21	Key words	X		X	Considered important therefore included
22	Language		X		Not included
23	Location	X	X	X	
24	Mood board	X			Not included
25	Office	X		X	
26	Professional/non-pro/UGC	X			Considered important therefore included
27	Project name	X			Considered important therefore included
28	Project number	X			Considered important therefore included
29	Project sector	X			Considered important therefore included
30	Project value	X			Considered important therefore included
31	Publisher		X		Not included – deemed inapplicable
32	Related press articles	X	X		
33	Resolution	X		X	Considered important therefore included
34	Source of image	X	X	X	Considered important therefore included
35	Subject		X		Not included – built into taxonomy and title
36	Status		X		Considered important therefore included
37	Staff name	X			Not included
38	Staff rank/position			X	Not included
39	Title	X	X		Considered important therefore included
40	Type of video	X			Not included – related to content type
41	Use (of structure)	X			Considered important therefore included
42	Value	X			Considered important therefore included
43	Version	X			Considered important therefore included

taxonomies and therefore was not required in the metadata. Elements of the subject would also be reflected in the title and file name. 'Type of video' was excluded, and it was deemed to be reflected within the content-type attribute. A more difficult exclusion, however, was the 'identifier' attribute. Although the identifier was considered important in uniquely naming content, it was also considered 'non-descriptive' and would not be used to either manage or retrieve content. Thus, it was excluded not on the basis of its usefulness, which was fully acknowledged, but on that of its 'non-descriptive' function. While excluded here, a conscious note was made to ensure that automatic software-generated identifiers would be used within the system when implemented. With respect to similarities, 15 of the 16 attributes identified in the desk study matched the user requirements highlighted in the workshops, whereas 13 of the 17 attributes from the standards reviewed also matched the attributes identified during the workshop. Thus, (without prior knowledge of the two activities) the outcome of the workshop was considerably similar to the outcome of both studies. The result of this refinement was a set of 35 attributes that reflected Company A's requirements, as shown in Table 3.

IMPLEMENTATION

RATIONALIZING THE ATTRIBUTES

In validating the appropriateness of the standard developed above, a total of six key potential users outside of the project team was consulted for a review. The review process consisted of issuing each person with the standard list and asking, on the basis of their role and the needs of their disciplines, which of the attributes should be deployed. The responses from the reviewers suggested that despite being comprehensive, the number of attributes was excessive and that it would prove to be a challenge for users, ultimately impacting the usability of the system. Also observed was that the integration of the proposed multimedia library into the wider IT infrastructure of the company would allow some attributes to be captured either automatically by the system (from the content itself) or as an existing attribute that could be drawn from other IT systems within the corporate enterprise architecture. Thus, a further refinement was carried out by the project team to regroup the 35 attributes defined on the basis of their importance, splitting the standard list into three tiers:

1 *Tier one*: Core attributes considered to be essential and that therefore should be implemented immediately with the system. This also included important attributes required only for administrative purposes, and not particularly descriptive metadata.
2 *Tier two*: Attributes required, largely to enhance the overall quality of metadata in the system and that are not critical to the effectiveness of the system.
3 *Tier three*: Attributes that, though required, add minimal value for the system.

Based on the levels of importance and to ensure the simplicity of use, it was decided that only tier one attributes would be deployed at the outset, whereas tiers two and three would be long-term additions, to be deployed as the system attained greater maturity. Thus the standard list of 35 attributes was reduced to an initial list of 20, as shown in Table 4. These were then further analysed to identify those that could be automated and those that would have to be manually input into the system by users. As for each of the attributes that could be automated, the team also identified its source (i.e. the internal system from which the data would be drawn). The outcome of this and of the deployed metadata standard is presented in Table 4.

LESSONS LEARNT

Having developed the metadata standard, a workshop was held with all team members to reflect on the activities carried out, the result obtained and to identify the lessons learnt. The key lessons were then captured by the researcher, analysed and summarized below.

IDENTIFY STAKEHOLDERS AND THEIR NEEDS

As it is crucial to ensure that the metadata used within the library meets the exact needs of the potential end users, a preliminary review is

TABLE 3 Thirty-five Metadata that form the standard for Company A

	ATTRIBUTE	DESCRIPTION
1	Access rights	Information on who can access the resource and associated details of the documents security level
2	Bid outcome	(For bid content) – An indicator specifying if the bid for which the content was created and/or used was successful
3	Business region	The specific company business region in which the office identified above belongs
4	Company	The company reflected in the theme of the content affiliated with or responsible for producing the content
5	Content type	The nature or genre of the content
6	Contributor (of image)	The entity responsible for providing the content to the library
7	Copyright	The nature of rights held in and over the content
8	Creator	An entity primarily responsible for making the content
9	Date of event/activity/ project	Date at which the theme or subject matter captured within the content occurred
10	Date image taken	Date at which the content was created
11	Date of commissioning (or use)	The date at which the entity in the content was commissioned
12	Description	An account of the content of the resource (ISO 15836, 2009)
13	Drawing no	(for CAD drawings saved as multimedia content) Unique identifier attached to a drawing
14	Designer (architect)	The entity/individual primarily responsible for creating the subject matter of the content
15	Duration/coverage	The extent of scope of the content (e.g. time of video and size of images)
16	File format	The digital manifestation of the content
17	Groups/disciplines	The company discipline/group reflected in affiliated to or primarily responsible for the subject matter reflected in the content
18	Interesting space	An indicator rating the innovative nature of the subject matter in the content. Typically reserved for buildings, structures and project content.
19	Key company staff	Key individuals within the organization affiliated to the subject of the content
20	Key words	Words describing the document's content
21	Location	The geographical position of the structure/event/entity captured in the image/video
22	Office	The specific office (of the company) reflected in, affiliated to or primarily responsible for the subject matter reflected in the content
23	Professional/non-pro/ UGC	The technical status of the contributor of the image
24	Project name	The name by which the project is formally known
25	Project number	The unique numerical identifier of the project affiliated with the image
26	Project sector	The specific work sector to which the subject matter of the content belongs
27	Project value	The total construction value of the project affiliated with the image
28	Relation	A link to any internally written press articles related to the subject matter of the content
29	Resolution	The density of the image expressed in words (e.g. high resolution, medium resolution, etc.)
30	Source of image	(If different from contributor) the resource from which the content is derived
31	Status	The state of the content in a related lifecycle
32	Title	Name by which the content is formally known
33	Use (of structure)	The use to which the subject matter of the content is/was subjected
34	Value	A rating system showing the relative significance of the subject matter within the content or the content itself to the company
35	Version	The current or previous states of the content

TABLE 4 Core metadata for multimedia library

	TIERS 1 AND 2 (CORE METADATA)	NOTE
1	Title	Manual
2	Content type	Manual
3	Description	Manual
4	Key words	Manual
5	Project number	Manual
6	Designer	Manual
7	Project phase	Manual
8	Access rights	Manual
9	Copyright owner/ notifications	Manual
10	Project name (auto)	Auto-generated from internal systems
11	Groups/disciplines (auto)	Auto-generated from internal systems
12	Location (country and city) (auto)	Auto-generated from internal systems
13	Sector (auto)	Auto-generated from internal systems
14	Creator	Auto-generated from content
15	Date image taken	Auto-generated from content
16	Duration for (video) coverage (for images)	Auto-generated from content
17	Contributor (of image)	Auto-generated from content
18	File format	Auto-generated from content
19	Resolution	Auto-generated from content
20	Version	Auto-generated from content

required at the outset of the project to fully define its scope, asking:

1 Who (in specific role/discipline terms) are the principal end users of the library? And who will be secondary end users?
2 What purpose would all end users be using the library for?
3 What metadata will be required to meet the specific needs highlighted above?
4 Of these metadata, what are the most important and which must be attached to all content?
5 What external regulatory or institutional requirements must be adhered to in defining this standard?

This stepped questioning process will result in a consistent definition of the project context that should be continuously referred to, in order to ascertain whether the elements defined are sufficient or excessive, or indeed unnecessary for the purpose intended. The project team observed that such a stepped questioning process was not properly followed at the outset, resulting in a distinct lack of focus in the type of metadata attributes required. For example, at the outset of the project, employees who would not use the library were extensively consulted and provided broad requirements that conflicted with the aims of the project (e.g. attributes such as the drawing-issue number, CAD-layering standards, etc. were requested, but these would not be required as the multimedia library would not be used to manage drawing files). Thus, refocusing on the core end-user base ensured that such attributes were not included and that the needs of the core end users were clearly focused upon. Drawing lessons from this, it is recommended that a focused stakeholder analysis be carried out right at the outset. It also promotes inclusiveness and enhances user buy-in for the project.

SIMPLICITY IS KEY

Participants unanimously identified the continual simplification of the number of attributes and of the terminology employed to be critical in ensuring that the metadata standard is both practical and pertinent, with one participant remarking that 'if at first look the system does not appear clear and straightforward, myself and most other users would simply create a folder on my computer and store the images there'. Such simplicity should be reflected in the terminologies chosen to define the attributes, the number of attributes selected, the taxonomy design, the design of the interface, the automation of certain attributes (to ease user input) and the training provided. Participants also reflected that it was because of the team's resolve to ensure the solution developed was simple and yet effective to the end user that the 35 attributes initially identified were refined even further to obtain a smaller subset.

METADATA STANDARD MUST NOT BE TECHNOLOGY DRIVEN

To ensure its usability, the metadata standard developed should drive the system requirements and not vice versa. Participants observed that

grounding the evolution of the metadata standard in the requirements of the users and their respective business processes rather than in the specific technological platform will serve to ensure that the solution is truly reflective of the end-user's needs. It will also ensure that the resultant standard aligns with the overall strategic IT architecture of the company and that it is not dictated by any technology. Although it is too early to assess the success of the standard developed, it was anticipated that a technologically agnostic standard, such as that developed here, will enable interoperability between internal systems and facilitate true company-wide collaboration.

STANDARDS SHOULD BE USED WHEREVER POSSIBLE

Participants observed that, considering the strength in similarities of the outcome from the workshop and the review of existing standards, the project should have begun simply with the review (or indeed with the adoption of an existing standard), using it as a baseline to build subsequent refinements. This, while providing a good starting point, will also ease information exchange between systems within the organization (and even external systems), as they would all be built to reflect similar metadata. Reflecting on the future, it was also noted that standardization would also ease migration of content to any future platform without a significant loss of metadata.

CUSTOMIZATION MAY ALWAYS BE NEEDED

Despite the adoption of external standards, customization may always be required to meet specific company needs and reflect specific business processes unique to each company (a point also raised by Paivarinta et al., 2002; CEN, 2005; Perlin, 2006; TASI, 2006). In further explaining this point, participants noted the difficulty in balancing the extent of customization to be carried out, as extensive customization may be time-consuming and provide limited value. While no clear means was identified to prevent this, a review of the case study showed that a second refinement was necessary to reduce the large number of attributes initially defined into a more focused set (a clear side effect of the extensive customization carried out).

STEPS TO DEVELOPING ORGANIZATIONAL METADATA STANDARDS

On the basis of the outcome of this case study, the specific processes undertaken and the lessons learnt (identified in the project closeout workshops), a guide is proposed depicting a process methodology for developing an organizational metadata standard. This is aimed at providing a usable framework with which Company A and other similar organizations could develop metadata standards to meet their needs. While based entirely on the outcome of the case study, the guide was also designed to reflect the following key characteristics (adopted from NAA, 2001):

1 Generic and flexible enough to be adopted to suit any organization looking to develop a metadata standard without being vague.
2 Multi-phased, focusing on steps to undertake and why those steps are essential. It makes no mention of specific tools to adopt but allows each organization to select the most appropriate tools to meet its needs.
3 Systematic, providing an easily workable structure.
4 Cyclical, accommodating a process of evolution and continuous redevelopment based on a whole lifecycle concept.
5 User-centric and focused on the needs of end users and business activities across the organization, thus building up from the overall objectives towards a workable solution.
6 Principles of project management including planning, resource management and change management are also reflected in the various stages of each process, emphasizing that the framework signifies an activity for which good management is deemed essential.
7 Compatibility with broader frameworks, such as the DIRKS model, that govern wider information and records management processes of which metadata development is only a part.

A total of 12 steps is proposed, grouped into three key phases as below:

1 *Establish goal of the project:* The first stage is to clearly define and articulate the intended goals of the project. A clear understanding ensures the right steps are taken and appropriate resources are made available for the effective execution of the project. This scope matched with the goal of the project enables strategic needs analysis to be carried out eliciting specific answers to the questions:
   a. Why should this project be carried out?
   b. What is the desired outcome of this project?
   c. What is the scope of the project?
   d. What business streams across the company will be affected by the project?
   e. What specific content classes are intended to be managed?
   f. Who are the target end users?
2 *Project initiation:* Here the goals identified above are translated into a working plan to articulate how the project will be managed from inception until completion. The business case and communication plan are developed including detailed justifications as to the viability and cost implication of the project. A high-level champion will need to be appointed to ensure top level support across the organization. The project delivery team will then need to be set up with a clearly defined mandate and responsibilities. In appointing the team (depending on the scope of the metadata standard being developed), it is important to ensure that membership cuts across the functional breadth of the organization so as to reflect the distribution of end users across the company and gain their input. The resources available to the team monetary and otherwise through the lifecycle of the project should also be defined along with a timeframe for its execution.
3 *Identify and analyse all related business processes:* Having initiated the project, a detailed analysis will then need to be carried out of the specific activities undergone by end users in carrying out those business processes. This stage is crucial as metadata do not exist in a vacuum. Its purpose is to contextualize a given activity to support the discovery and management of content to support such activities. Hence this stage aims to identify what those activities are for which the metadata is required from the end user perspective. The outcome of this would also be used to validate the eventual standard developed to ensure it meets the intended goals. Thus, this stage answers the questions
   a. What are the specific business processes for which the metadata is required?
   b. Within each of those processes, what activities are carried out which require metadata?
4 *Identify metadata requirements:* Based on the understanding of the activities and process, this stage then identifies the nature of metadata required to achieve it. It is important to maintain the order starting with an understanding of the processes and from that building up a picture of the sort of metadata that would be needed to support it. At this stage no details are required of the individual metadata attributes. What are required are themes such as: subject metadata, administrative metadata, regulatory metadata, retrieval metadata, workflow metadata, etc.

These four steps constitute the project definition phase of the project, giving it the appropriate direction required to execute the other phases. A clear definition is required at the end of these four steps to ensure that the steps within the subsequent phases are both precise and adequate. The next phase is the metadata development phase.

5 *Review existing standards:* Standards should then be reviewed to identify which could be adopted to meet the needs of the company and its business processes. A broad range of standards should be consulted and selected based on their internationalization, perceived relevance to the organization, relevance to the industry in which the organization is based and/or relevance to the type of content being managed. While these selection criteria may vary across companies and indeed projects, they need to be clarified to ensure the right sets of standards appropriate for the project are reviewed. Some key questions that should be answered here include:
   a. What criteria should be used to determine what standards are appropriate for the company's needs?

b. What international/cross industry standards are available which meet these criteria?
c. What industry-specific standards are available which meet these criteria?
d. What content-specific standards are available which meet these criteria?

6 *Review appropriateness:* Having reviewed and identified potentially suitable standards, these then need to be analysed to determine their suitability for the needs of the project. This involves a detailed analysis of the individual attributes contained within each of the shortlisted standards above (as similarly carried out in the case study presented). Where a standard is deemed wholly appropriate and can be adopted with little or no customization this can then be carried on to the testing stage. Where customization is deemed necessary, this activity saves time and resources by identifying those attributes that can be adopted as a base minimum, serving as a good starting point from which the required customization can then be carried out. The detailed metadata elements agreed upon are also refined here, identifying variations such as compulsory and optional attributes; or automatic and manual attributes. The success of this stage requires the crucial input of end users to confirm the suitability or necessity of any of the attributes. Where customization is carried out, the review process should be done iteratively until agreed as being appropriate.

7 *Customize standards:* Where no standard is appropriate, modification is carried out at this stage to make the standard reflective of the needs of the company. It should begin with a clear understanding as to why the modification was deemed necessary. The result is a company-specific metadata standard that takes into account such varying perspectives as business processes, archiving policies, quality management procedures, business structure and so on. Each attribute identified here must be justified, identifying why it is necessary and by whom. This is to ensure that while additional attributes are included, the standard developed remains fit-for-purpose. Various methodologies can be used for capturing the elements required here including desk studies of existing repositories, workshops, questionnaires and so on. Irrespective of the methodology, the eventual outcome must be collectively reviewed by the project team accepted and signed off as appropriate prior to any testing or validation.

8 *Testing and validation:* To ensure its suitability, the metadata standard should then be tested around various scenarios within its anticipated scope of use. These scenarios should be as varied as possible. The object of this exercise is to scrutinize the standard for any loopholes and ensure the solution is robust enough to meet the needs of the company and its wider user community (see Figure 1).

These represent the four steps to be carried out at the metadata development phase. This phase where details such as the individual elements and their associated definitions are developed represents the translation of the project vision into a workable standard. It should be noted that the case study discusses work carried out up to this stage (and does not include details of any actual implementation). Thus the next steps outlined below are anticipated steps as defined by the project workshops as well as findings from detailed literature reviews. They are currently being validated through the on-going implementation process.

9 *Develop governance approach:* Governance enables the distribution of accountability and responsibility for the long-term management of the metadata standard. This is necessary in metadata development to ensure that it is continually updated to meet the future needs of users and the company in light of changing business strategies, methods of working, processes and regulations (Bentley, 2001; Sun Microsystems, 2005). Specific decisions will need to be made including (but not exclusively): how the standard will be managed in the future, who retains responsibility, how the quality of metadata input into the system can be ensured, etc. (Paivarinta and Munkvold, 2005). As responsibilities are assigned, resources will have to be committed for this purpose. Change management is a significant activity carried out at this stage aimed at

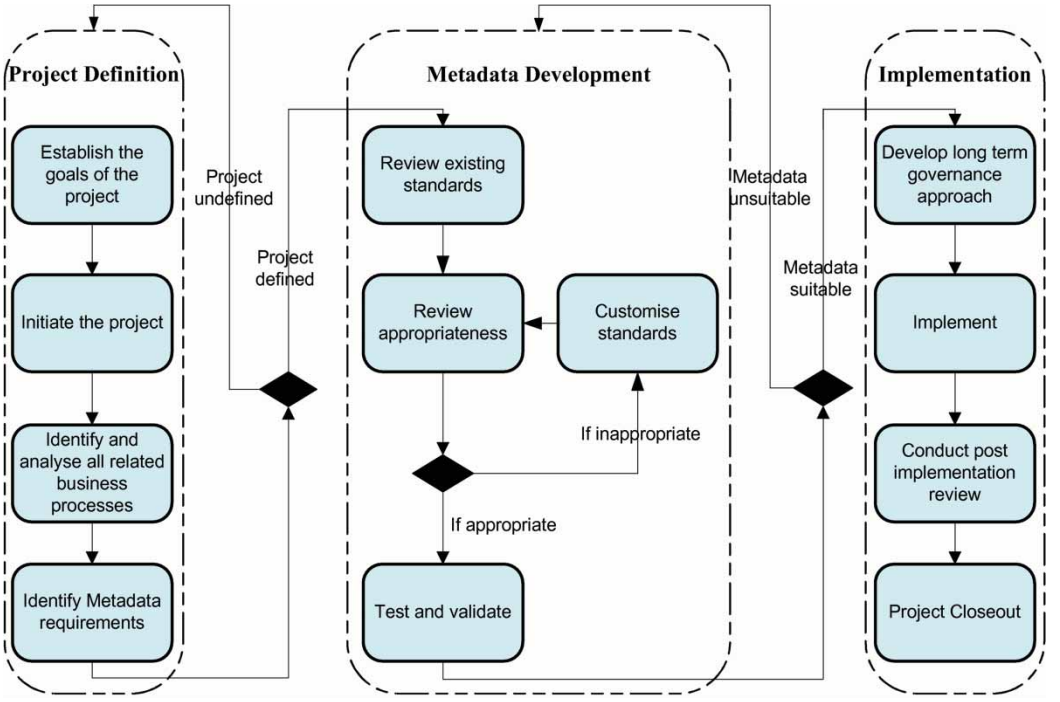

FIGURE 1 Steps for developing an organizational metadata standard

facilitating a smooth transition to the adoption of the developed standard. It should be seen as a conscious activity required to ease user adoption and could significantly impact on the success of the project (Mathieu and Capozzolli, 2002).

10 *Implementation:* This is the stage in which the developed standard is encoded into the proposed system. The reference here is to the visual interfaces and the actual representation of the metadata on the chosen system as seen by the user; and interfaces for both metadata entry and result visualization including critical questions on: which of the attributes should be free text or automated. Beyond the technological implementation, necessary training and policy guidance developed earlier are also rolled out here in a suitable manner.

11 *Post-implementation review:* Having implemented the developed standard, the processes undergone and the outcome of the project should then be collectively reviewed. This helps to ascertain if the original project goals were met. This reflection also enables lessons learnt to be recorded to improve the delivery of any future metadata development endeavour. Also required here is feedback from end users to establish the appropriateness of the metadata standard developed to meet their needs.

12 *Project closeout:* This marks the formal end of the project and is carried out when the metadata standard is deemed functional and fully implemented. This also marks the beginning of the long-term maintenance of the standard (as defined when developing the governance approach).

DISCUSSION

The standard developed through this exercise for Company A though specific to multi-media content, would form part of the broader metadata standards (to include all other unstructured content) also being currently developed. Looking through the 35 attributes, a significant observation is that many

of the attributes are not unique to Company A, but are also similarly applicable to other organizations. Thus, further analysis was carried out to investigate the similarities between the attributes based on their varying levels of standardization. Accordingly, four significant groupings emerged. These were:

1 *Generic standard metadata* – Attributes reflected in existing cross industry standards. A total of 16 of these were identified. Example of such a standard is the ISO15836.
2 *Content-dependent metadata* – Attributes required due to the nature of the content being managed within the system (in this case multimedia). These could be standardized across content classes. Content classes such as drawings, models, etc. can cut across companies and industries. Three attributes were identified here. An example of a content dependent standard is ISO 13567 (2D Computer Aided Design standards).
3 *Industry-dependent metadata* – Attributes required due to peculiarities in the industry/sector in which the company is based (in this case the construction industry). Two attributes were identified here. An example of such a standard is the BS1192:2007.
4 *Custom metadata* – Attributes that reflect the specific needs of the company with all its processes, workflows, activities and quality management procedures. Fourteen attributes were identified here.

These groupings suggest that organizational metadata for all companies across industries cannot be standardized without some degree of customization to suit particular company needs. This also sheds light on the challenge raised by Paivarinta et al. (2002) of where the appropriate starting point is for defining organizational metadata. With 20 of the 35 attributes (and 13 of the 20 tier one attributes) in this study based on standards defined outside the context of the specific organization, these results suggest that building up from existing standards may be a better starting point than developing wholly bespoke solutions. In this case, such an approach would have provided over 50% of the metadata required. Owing to the focused case study approach adopted for this research, no attempt is made at generalizing these findings as further research is still required to ascertain the universal validity of the results.

The 12 steps outlined above are proposed based on a retrospective analysis of the outcome of the project (from the workshop as well as a detailed process analysis conducted by the researcher and project team). Thus, the project in its execution did not entirely adhere to the steps above but underwent iterative processes with significant challenges. While the goals of the project were clearly defined at the outset, the scope of who the target end users would be was not. Thus an immediate challenge was clarifying the scope of the type of attributes that would be most appropriate to the target group. It was observed that strictly clarifying the target groups at the beginning, as being marketing and business development, would have helped ensure that the standard that was focused to reflect their business processes and needs.

The lack of extensive experience in the process of defining metadata within the organization also meant that the project initiation process was a challenge. There was an initial lack of understanding of the true value of metadata and therefore the necessity of the tasks to be carried out. Thus the project manager begun first of all by enlightening the entire team of the necessity of the endeavour, the magnitude of work that will be required and the thinking required to execute it. Also, the team was assembled from across the company to include a project manager and representatives from all the user groups with the researcher providing guidance on metadata and taxonomy. During the closeout workshop it was observed that the cross-disciplinary membership of the project team, reflecting the actual end user-groups, ensured clarity in the requirements and the way the team approached all tasks. It also helped ensure that all the related business processes that the content library was procured to support were consistently reflected upon and referred to through the development process.

Another challenge faced was in reviewing the appropriateness of the developed standard. To ensure that fresh perspectives were incorporated, a total of six individual users outside the project team were consulted. Having reviewed the initial 35 attributes proposed, all six remarked that the list was 'too long' and included attributes that they (individually and their respective disciplines) would not use. However, when asked to specify their preferred attributes, all selected attributes from within the list and no new attribute was included thus confirming the comprehensiveness of the standard. Thus, the problem as they saw it was not as much the standard not being reflective of their needs, but that as it also reflected the needs of other staff outside of their business streams (it appeared to be much broader than they anticipated). Also significant was that as the six individuals belonged to different business streams, there were variations in their selections with some arguing for the importance of certain attributes to support their tasks and others seeing those attributes to be unimportant. The core 20 attributes to form the tier one attributes were generally accepted by the reviewers as being sufficient, with the understanding that, with the maturity of the system, these could be expanded to include the other 15 attributes that make up the standard.

While the review described above helped in refining the attributes and ensure its appropriateness, there was still a need to test whether or not the attributes and the standard developed actually added value to end users when the system was fully deployed. This implementation process involving rolling out the software to the organization is currently in progress. Thus, testing the appropriateness of the developed standard is to be carried out as part of the next phase of the research. Because the aim of using metadata was to primarily facilitate information retrieval, the appropriateness of whether the standard is 'good' or not will be measured by whether:

- content is more easily retrieved from a system that uses the attributes (as compared to a system without the attributes);
- content is better organized in a system that uses the attributes than in a system without it; and
- were the metadata truly reflective of the needs of the end-user community? That is are these the right attributes? Are there any attributes required that are not captured?

LIMITATIONS OF THE STUDY

The focused cased study methodology adopted here enabled the exercise to be carried out in detail. However, it implies that the conclusions from this work while contributing to knowledge cannot be at this point generalized. Further research will need to be carried out along with other case studies to test the viability of the findings and the robustness of the framework in various contexts. The project and the metadata standards presented here are currently being implemented within the content management system, therefore the findings of this study do not include a practical assessment of the actual implementation of the standard, the user response to this and its impact on information retrieval within the organization. While validation was carried out with some end users, a practical assessment of the standard after its full implementation will be required to further validate the findings and the steps proposed. Also not assessed here is the impact the implementation process may have on the guide proposed.

The researcher and indeed the project team were constrained by limited resources and were therefore unable to engage with a larger sample of users for the validation of the attributes; hence only six users were consulted. It is anticipated that a survey will be built into the system when launched to capture user feedback so as to improve the standard as part of the long-term governance approach. As also explained, a workshop was conducted with the project team to review the outcome of the project. While this ensured collective perspectives were discussed and then captured, it also served to limit the identification of individual problems, concerns and issues faced by each of the user representatives. In the future, these would perhaps be better captured via unstructured interviews.

CONCLUSION

This article presented a case study in which a metadata standard for a proposed multi-media

library was designed with the detailed processes undergone. The three tasks, that is, desk study, review of existing standards and the workshop were explained along with their detailed findings. The resultant metadata standard developed was then presented with details of all the analysis carried out. Specific lessons learnt through the whole process were clearly outlined culminating with a proposed step by step guide for how organizations can develop similar metadata standards. As well as the case study presented, the findings add to knowledge by defining how the process of metadata creation can be more systematic. Other organizations looking to embark on similar endeavours can find within this study lessons to guide them through the process.

By exploring alternative approaches to developing metadata standards, this study also addresses the question of what an appropriate starting point is in developing a metadata standard. The strong similarities between the baseline developed and the ISO 15836 (2009) standard showed that using such standards would be a suitable starting point upon which to build an appropriate standard. Similarly, the findings support the thesis that no specific standard is comprehensive enough to meet the needs of an organization without appropriate levels of customization. Such customization should be built entirely around user needs and the processes which the solution is designed to support. While this exercise focused on multi-media content, it is anticipated that within Company A it will be used to develop a wider metadata standard for all unstructured content (i.e. documents, drawings, models, simulations, etc.) to ensure consistency. Further research is needed to test and determine the practicality and suitability of the proposed guide for use by other companies within and outside the construction industry and also for managing other types of content beyond multimedia files.

REFERENCES

Archives New Zealand, 2004, *New Zealand Government Locator Service, Version 2.1, (online) New Zealand* [available at www.e.govt.nz/standards/nzgls/standard/element-set-21/] (last accessed, 18 January 2009).

Bentley, J.E., 2001, *White Paper on Metadata: Everyone Talks About it, But What is it?* (*Online*) [available at www.sas.com] (last accessed 25 February 2008).

Bjork, B.C., 2001, 'Document management- a key IT technology for the construction industry', *ECCE Symposium*, European Council of Civil Engineers, Finland.

Bock, G., 2005, *White Paper on Designing Metadata – An Implementer's Guide for Organising and Using Digital Assets* (Online) [available at www.databasics.com.au/docs/Designing_Metadata.pdf] (last accessed 18 May 2009).

British Standards Institution, 2007, *BS 1192:2007, Collaborative Production of Architectural, Engineering and Construction Information – Code of Practice*, BSI, Milton Keynes.

Burnett, K., Ng, K.B., Park, S., 1999, 'A comparison of the two traditions of Metadata development', *Journal of the American Society of Information Science* **50**(13), 1209–1217.

Cabinet Office, 2004, *UK GEMINI Standard – A Geospatial Metadata Interoperability Initiative*, UK.

Day, M., 2006, 'Metadata – a general introduction, presentation from cataloguing online resources conference', *Manchester, UK*.

E-Government unit, 2006, *E-Government Metadata Standard Version 3.1*, Cabinet office, UK.

European Committee for Standardization (CEN), 2005, *Guidance Information for the Deployment of Dublin Core Metadata in Corporate Elements*, Workshop Agreement, CWA 15247, Brussels, Belgium.

Haynes, 2004, *Metadata for Information Management and Retrieval*, Facet Publishing, London, UK.

Hofman, H., 2006, Standards: not 'One Size Fits All', *Information Management Journal*, **40**(3), 36.

Institute of Electrical and Electronics Engineers, 2002, *IEEE 1484.12.1 – Standard for Learning Object Metadata*, New Jersey, USA.

International Standards Office, 2009, *ISO 15836 – Information and Documentation – The Dublin Core Metadata Element Set*, ISO, Geneva.

International Standards Office, 2001, *ISO 15489 -1 – Information and Documentation – Records Management. General*, ISO, Geneva.

International Standards Office, 2003, *ISO 19115 – Geographic Information – Metadata*, ISO, Geneva.

Karjalainen, A., Paivarinta, T., Tyrvainen, P., Rajala, J., 2000, 'Genre based metadata for enterprise document management', *Proceedings of the International Conference on System Sciences*, *Hawaii*.

Mathieu, M.L., Capozzolli, E.A., 2002. 'The paperless office: accepting digitized Data', *System-wide Business Symposium*, Troy State University.

Murphy, L.D., 2001, 'Addressing the metadata gap: ad-hoc digital documents in organisations', in: A.G. Chin (ed), *Text Databases & Document Management: Theory and Practice*, Idea Group Publishing Hershey, PA, 52–57.

National information standards organisation NISO, 2004, *Understanding Metadata*, NISO press, Bethesda, MD, USA.

Paivarinta, T., Munkvold, B.E., 2005, 'Enterprise content management: An integrated perspective on information management', *Proceedings of the 38th International Conference on System Sciences*, *Hawaii, USA*.

Paivarinta, T., Tyrvainen, P., Ylimaki, T., 2002, 'Defining organisational document metadata: a case beyond standards', *Proceedings from the European Conference on Information Systems*, ECIS, June 6–8, Poland, 1154–1163.

Perlin, N., 2006, 'Introduction to metadata', *Proceedings of the International Professional Communication Conference*, IEEE, 23–25 October, New York, USA, 153–155.

Rockley, A., Kostur, P., Manning, S., 2003, *Managing Enterprise Content, a Unified Content Strategy*, 1st edn, New riders, California, USA.

Sun Microsystems, 2005, *White Paper on Metadata Management: An Essential Ingredient for Information Lifecycle Management* (Online) [available at www.sun.com] (last accessed 25 November 2008).

Technical Advisory Service for Images, TASI, 2006, *Advice Paper on Metadata Standards and Interoperability* (Online) [available at www.tasi.ac.uk] (last accessed 25 February 2008).

The National Archives, 2002, *Requirements for Electronic Records Management Systems: 2 – Metadata Standard*, Surrey, UK.

The National Archives of Australia, 2001, 'DIRKS – a strategic approach to managing business information', *Part 1 – The DIRKS Methodology – A User's Guide* (Online), Australia [available at www.naa.gov.au/records-management/publications/dirks-manual.aspx] (last accessed 15 May 2009).

The National Archives of Australia, 2006, *Australian Government Implementation Manual: AGLS Metadata*, Version 2.0 (Online), Australia [available at www.naa.gov.au/records-management/publications/AGLS-manual.aspx] (last accessed 15 May 2009).

White, M., 2005, *The Content Management Handbook*, Facet Publishing, London, UK.

Yin, R.K., 2002, *Case Study Research – Design and Methods*, 3rd edn, Sage Publications Inc., Newbury Park, CA.