Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia - Social and Behavioral Sciences 73 (2013) 551 – 555

Procedia
Social and Behavioral Sciences

The 2nd International Conference on Integrated Information

Dynamic indexation in video metadata

Camila Ribeiro[a]*, Marcos L. Mucheroni[b]

[a]*University of São Paulo. Department of Postgraduate Program in Information Science, São Paulo, Brazil cribeiro@usp.br*
University of São Paulo. Department of Postgraduate Program in Information Science, São Paulo, Brazil mmucheroni@usp.br

Abstract

This paper is an exploratory study about metadata for video database. It has as purpose establish a dynamic indexation methodology for multimedia video environment. Thereafter the popular models of textual publication, for instance the OJS, have popularized Dublin Core as representation pattern. It already is broadly used in scientific papers dissemination, however even for text and images the analysis of "aboutness" and "offness" must be taken to enable dynamic indexation, which is essential for multimedia environments. This paper proposes an element expansion for Dublin Core to reach this goal.

Keywords: Dynamic indexation. Dubli Core. Dynamic information. Video

1. Introduction

The increase of information on the web has brought documents convergence, from printed to virtual, consequently the digital information on the web spreads in several types of media (music, video, images, animations, etc...) making obsolete the classic textual information retrieval techniques. As D'Andréa (2006) [1] says, the structural changes in document environment caused by documents digitalization and by virtual communication have brought definitive impacts on configuration of documents even in its internal structure and in operational logic. The virtual forms are revolutionizing the concept of documents [2].

If textual information retrieval on the web is becoming harder due informational volume, on the other hand non-textual digital information has not a solid representation technique, even if some people are working to develop this technological lack [3].

* Corresponding author.
E-mail address: author@institute.xxx

doi:10.1016/j.sbspro.2013.02.090

In libraries ambit traditionally the information availability is given by representations, named since the 90's "metadata"[4]. according to Souza and Alvarenga (2004) [5], "Meta" is a self-reference prefix therefore "metadata" means "data about data", in a sense of data retrieval. Metadata in web documents has as function specify data characteristics such as ways of use and ways to display. It can also specify the meaning in certain context recovering the semantic nature of it, which is frequently an aspect of documentation field research.

Thus, the metadata study is related with rationalization knowledge process by which the world's society is passing during these last years. The models and formats to metadata representation that exist, such as Dublin Core, MARC (*Machine-Readable Cataloging*), RDA (*Resource Description and Access*), inter alia, are insufficient or incompatible [6]. First of all, the adoption of formats must be analyzed in greater detail to subsequently find out the real needs of each media, searching for available software and its possibilities. When becomes necessary a richer description, it emerges also the need of more expressive data and vocabulary models to define and delimitate a broader scope [7]. Based on that need it was created the OAI, a data model for scientific papers, built on DC (Dublin Core), but at the same time concerned with textual data representation.

In thematic representation of information, particularly in indexation, the document content named by Caffo (1988) [8] as "semantic document content" is, according the author, the object of a special way to access information. The proposition is a phase of conceptual analysis followed by the translation from analysis results to an indexing language. To the author, inherent to indexation process is "the document analysis to identify its semantic content as well as the translation of this content into certain documentary language" [9].

To Information Science the indexation of information in textual format is classic established although there are indexing differences for other formats such as image and video, etc. because these formats require a dynamic representation that fulfill the user's and author's aim. Often the object of an image is hard to identify clearly, so it is necessary give more attention to other access point in order to retrieve, hence is essential think how organize and express it carefully [10].

Indexation can be viewed under two perspectives: the "offness" – the concept meaning, relative to document semantics [11]; and the "aboutness" – what the document is about and its descriptive concept -, both together form the dynamic description, which concerns this paper.

Hence, to dynamic indexation of videos is not possible use the same search engines of textual contents in which the context has a formal structured thinking. Zhong (2001) [12] brought to light an important aspect about image treatment, segmentation, analysis and abstraction. Regarding the subject access during indexation Berinstein (2003) [13] reminds that context, relationship, "historical significance" and other factors must be taken account of and somehow represented by the visual arts cataloger. Eakins and Graham (1999) [14], give an overview of the complexities of image cataloging, suggesting that the needs of users are central to determining focus and depth of indexing [15].

In the literature of video indexing many heuristic methods are proposed. The most advanced techniques explicitly use pattern recognition. The improvement of information sources access by the users, mainly due images and videos, is becoming increasingly heterogeneous [16]. Given the circumstances we choose the XML language.

Documentation for specific video snippets depends on two simultaneous processes: video segmentation (there are already computation techniques regarding this); and the definition of documentary formats using tags or metadata. There are few segmentation models, such as W3C SMILL (Synchronized Multimedia Integration Language), DC (Core Metadata Element Set Dublin), Media RSS, and several metadata formats for internet files, such as MARC and FRBR. For this research we choose Dublin Core by being the most popular format between academics. It is simple and already has a semantic structure (XMLSchema), which is a defined field structure allowing to set parameters of "aboutness" and "offness", so the video documentation is dynamic as in Kacher *et al* (2003) [17], except in this case is used for photographs, that are also static.

[...] "OFFNESS ". It is an approach related to the extractive information from the picture. We identify: • The graphic content ; concerns everything directly legible on the picture (colour, texture, shape,...) • The semantic

content ; concerns the identification elements represented and legible directly on the picture: The other family of describers are the so-called "ABOUTNESS". It is an approach that carries on the complementary or the interpretative but no directly extractive information from the picture. We identify: The contextual content ; is related to concepts no legible directly but having a obvious connection with the content of the picture (type of project, name of project, geographical situation,...) • The symbolic content is located on a more elevated abstraction level which is the semantic level (opacity, permeability, monotony,...).

The dynamic indexing proposal includes two types of image, and creates a connection so-called “relation between the parts”. On the DC we used the relation hasPart element, as visualized in the picture.

DC traditional symbolic elements as title, creator, description, date, etc, can be founded as well. The dynamic links can be established by DC element relation hasPart, and the “offness” is inserted on the video elements: sequence, scene, shot, frames. All these must have their specifications.

So the RDFS description that establishes the application profile creates a dynamic structure to describe videos.

2. Methods

2.1. Proposal using Dublin Core

Dublin Core was designed specifically for generating metadata for textual documents. Although a number of workshops have been held to discuss the applicability of Dublin Core to non-textual documents, they have primarily focused on extensions to describe bibliographic-type information rather than the actual content [18]. The authors presented the findings and discussions during a workshop, where they concluded that DC potential use to describe moving images is a solution to the problems, with some caveats.

The definition of proposed fields for video documentation has to be in RDFS format, which allows dynamic video description and allows to set a group of formats in an application profile. The proposed fields must be in the semantic structure in the XML specification. That makes easier the textual research and the semantic control by intelligent agents [19].

XML was designed to be a simple way to send documents across the Web. It allows anyone to design their own document format and then write a document in that format. These document formats can include markup to enhance the meaning of the document's content. This markup is "machine-readable," that is, programs can read and understand it [19].

Several organizations believe that XML will be the major standard enabling better interoperability between systems and users [20]. XML also offers an excellent adaptation support. By using XML, the content author can semantically mark up the contents of a document, describing the content in terms of its relevance as data. Software that processes XML does not need to know the semantics of the document. XML does not specify how the contents of a document should be presented or saved. However, there are several XML-based markup languages, e.g. XHTML, that have fixed semantics for layout to enable content representation with standard browsers [16].

Considering that the non-textual system is interoperable, a suggestion is a set of XML elements, for instance the 15 DC textual elements because it is used in many systems (such as OJS – Open Journal System). The XML elements and an explanatory figure 1 the flowing.

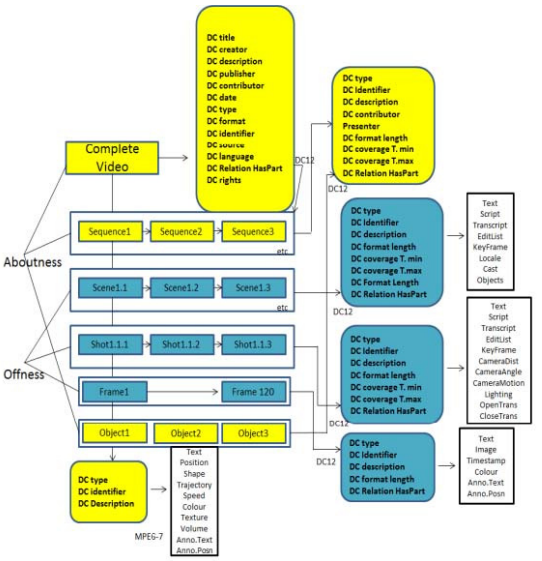

Fig. 1. Dublin Core applied to video.

Blue elements are “offness”, and they have information directly from the video. Yellow elements are “aboutness”, they are those that focus the appearance and have to be descripted dynamically.

Indexing basic unit is the scene (work as a word for a text document), and a group of scenes form a unit that is indexed in DC element 12, where is possible to have a dynamical semantic description. Scenes work like paragraphs in a text while the sequences work like pages or chapters. A group of scenes that can be divided is a segment. For instance: a TV video has 25-30 fps, each frame is an image, but if isolated it doesn’t represents anything besides itself, but when connected in sequences it must have more dynamic connections, composing scenes, sequences and even segments.

Video segmentation must obey a logical structure. Unlike texts, that are easily segmented in paragraphs and topics, video structure is different since you have to find self-contained snippets that can be indexed without speech or subject rupture.

3. Conclusion

To conclude is good remind that video data are particularly difficult to index due to the dynamism, and are not static as most of metadata patterns propose, hence these patterns are inadequate for videos, as previously discussed. Bibliographical systems have different video language expressions. In films, videos, TV shows, the contents are added or edited defining video segments as “another” reference on DC 12nd element (in the figure they are indicated as objects), but most of them are lacking in indexation and in copyright references.

The culminating difference of video metadata indexation is how the user sees content, i.e., scenes, lines and snippets that he wants to find and the current video description systems can’t handle.

References

[1] D’Andréa, C. “Estratégias de produção e organização de informações na web: conceitos para a análise de documentos na internet”. Ci. Inf., Brasília, 35, 39-44, (2006).

[2] Pédauque, R.T. “Document: forme, signe et médium, les reformulations du numérique”. Version 3, (2003).

[3] Galdo, A.; A.F.G Vieira and R.S. Rodrigues. “Classificação social da informação na web: tecnologia, informação e gente”. DataGramaZero, 8, (2009).
[4] Dziekaniak, G. “A Organização da informação e a comunicação científica: implicações para os profissionais e usuários da informação”. Em Questão, Porto Alegre, 16, 43-57, (2010).
[5] Souza, R.R.; L.A. Alvarenga. “Web Semântica e suas contribuições para a ciência da informação”. Ci. Inf., Brasília, 33, (2004).
[6] Siqueira I.C.P. and J.F.M. Silva. “Metadados: o fio de Ariadne ou a coragem de Teseu?” Bibl. Univ., Belo Horizonte, 1, 11-18, (2011).
[7] Lagoze, C. “Keeping Dublin Core simple cross-domain discovery or resource description?”, D-Lib Magazine, 7, (2001).
[8] Caffo, R. “Analisi e indicizzazione dei documenti”. Ed. Bibliografica, Milano, (1988).
[9] Guimarães, J.A.C. “Abordagens teóricas de tratamento temático da informação: catalogação de assunto, indexação e análise documental”. Ibersid, 105-117, (2009)
[10] Turner, J.M. “Indexing pictures: some considerations”. Annual Meeting Council on Botanical and Horticultural Libraries, Montréal, (1997).
[11] Fairthorne, R.A. “Content analysis, specification, and control”. Annual Review of Information Science and Technology, 4, 73-109, (1969).
[12] Zhong, D. “Segmentation, index and summarization of digital video content”. Columbia University, (2001).
[13] Berinstein, P. “Moving multimedia: the information value in imagens”. Searcher 5, 40-48, (1997)
[14] Eakins, J. and M. Graham. “Content-based imagem retrieval: a report to the JISC technologies”. Apllications programme, (1999).
[15] Barta-Norton, N.A. “Marc applications for description of visual materials”. J. of education media & libary sciences, 42, 21-36, (2004).
[16] Forstadius, J. and M. Löytynoja “XML in dynamic multimedia content management”. Proceedings of 2001 Nordic Interactive Conference, Copenhagen, (2001).
[17] Kacher, S., J.C. Bignon and G. Halin. “Image indexing vocabulary in architecture taxonomic hierarchy and categorisation”. E-Activities and Intelligent Support in Design and the Built Environment, the 9th EuropIA International Conference, Turkey (2003).
[18] Hunter, J. and R. Iannella. "The Application of Metadata Standards to Video Indexing" in C. Nikolaou and C. Stephanidis, eds. “Research and Advanced Technology for Digital Libraries: Second European Conference”, ECDL '98, Heraklion, Crete, Grece, September 1998, Proceedings. (Lecture Notes in Computer Science 1513) Berlin: Springer, (1998).
[19] Swatz, L. “Why people hate the paperclip: labels, appearance, behavior and social responses to user interface agents”. thesis to bachelor of Science, advisor: Clifford Nass, Stanford University, (2003).
[20] Miller, J.A. and S. Sheth “Querying XML documents”. IEEE Potentials, 19, 24 –26, (2000).