

## **MEDINA: A Semi-Automatic Dublin Core to MPEG-7 Converter for Collaboration and Knowledge Management in Multimedia Repositories**

**Marc Spaniol and Ralf Klamma**

(Lehrstuhl Informatik V, RWTH Aachen, Germany)

{mspaniol|klamma}@cs.rwth-aachen.de

**Abstract:** Knowledge creation processes in the cultural sciences have a discursive nature. The cultural scientists participating in these discourses build a community of practice. Reifications in information systems should support four concepts: (1) Content description by metadata, (2) coverage within standards, (3) repository technologies, and (4) platform independence of applications. By the combination of these concepts information systems support cultural scientists to extract and manage knowledge about high-level semantics of multimedia artifacts in open repositories with metadata annotations. Basic support for the latter can be achieved by a loose classification scheme as in Dublin Core, but with more sophisticated MPEG-7 description elements for time based media. We present a system called MEDINA for semi-automatic Dublin Core to MPEG-7 conversion to maintain media already annotated in Dublin Core.

**Key Words:** Knowledge Management, Communities, MPEG-7, Dublin Core, Information Systems

**Category:** H.3.0, H.5.1, J.5

### **1 Introduction**

The support of collaboration and knowledge management (KM) in networked cultural science communities represents a specific challenge for the development and the organizational structure of information systems. The reason is inherent in the specific research methods cultural scientists apply and their media centric way of collaboration that facilitates the overall discourse. In our collaborative research center “Media and Cultural Communication” we as computer scientists have analyzed the impact of networked information systems (IS) on the organization and creation of knowledge during the last six years. We have realized dedicated IS with them. Our experience shows that these systems have to be carefully designed with respect to at least four aspects.

First, the discursive nature of functioning in the humanities, requires support of different digital media to be combined with (almost) arbitrary metadata. Thus, concepts for collaboration and KM in the humanities require a high degree of semantic freedom. Second, metadata standards are needed to bridge the gap between loose classifications by the community and the structured organization of repositories. In this aspect, the Dublin Core (DC) metadata standard [5] has been a step forward. It’s a simple, light-weight and concise method for

resource discovery of composite resources on the web. However, this standard is not designed to describe temporal and media specific information connected with multimedia resources used in the communities we cooperate with. Thus, we have selected an excerpt of the extensive MPEG-7 multimedia metadata standard [10]. Hence, we can offer facilities for a detailed metadata annotation of media specific information as well as temporal decompositions. Third, a common repository is an indispensable prerequisite to foster knowledge exchange. An open community KM system requires to offer its users an option to not only access multimedia artifacts already in the system, but to contribute multimedia artifacts to the repository, too. Fourth, platform independence is an indispensable prerequisite. Although, MS Windows PCs are wide spread within the cultural sciences, particularly users in the film studies - we are cooperating with - commonly use Apple computers.

KM strategies thus need to bridge the gap between semantic freedom in (almost arbitrary) metadata annotations and (structured) repository technologies. In this paper, we present an application called MEDINA (MPEG-7 Encoding of Dublin Core Information and Naming Application) that allows the administration of open multimedia archives in film study communities. The rest is organized as follows. First, we introduce requirements for collaboration and KM in cultural science communities. Then we discuss existing multimedia management systems regarding their alignment to the needs of cultural science communities. After that, we present MEDINA to comply with the needs of our target community. The paper closes with conclusions and an outlook on further research.

## 2 Collaboration and KM in cultural science communities

The research done within the last six years of our project has given us a detailed insight into work practices of cultural scientists. Collaboration and KM in cultural science communities is coined by the discursive nature of functioning within the cultural sciences in general. Basically, there are two types of knowledge organization and exchange within scientific communities as introduced by Snow [18]. First, the ‘linear type’ of learning that is goal-oriented and transmission-centered. This means, old information is replaced by new one as soon as this appears. Second, there is a ‘non-linear type’ of KM. This type is media centric and reflects working practices of cultural scientists. It doesn’t replace old knowledge but keeps it in its repositories for latter usage. Thus, an ongoing description and (loose) classification of medial artifacts is necessary for them to bring their studies further. Based on the previous considerations, KM strategies for cultural scientists require a high degree of semantic freedom. That means, it is not sufficient to examine KM systems and technologies for multimedia annotation strictly separate as it is propagated in state-of-the-art KM systems. For

that reason we try to overcome their shortcomings as stand-alone solutions by combining them in an open and collaborative manner.

In order to foster the collaboration when working on multimedia artifact sets, first of all, an easy to use import (and export) functionality for already annotated media is needed. Due to the dispersion of DC annotations in print media, import options for annotated DC artifacts are eligible. Moreover, cultural scientists are particularly sensitive to easy to use user interfaces hiding any technical complexity. Additionally, authentication is needed in order to protect the intellectual property on information about media sets as well as the media sets themselves. In the following section we give an overview on related work dealing with multimedia management systems and discuss whether these systems comply with the aforementioned needs of cultural science communities.

### 3 MM management systems for cultural science communities

There are several tools and projects dealing with multimedia annotations. Most of them are situated in the field of video annotation and deal in particular with the fine grained analysis of videos. The resulting annotated multimedia artifacts are mostly compliant to one standard, either MPEG-7 or DC. Except some web-based projects most of the systems are restricted to MS Windows PCs. We now present related projects in detail.

Our overview on related work starts with metadata annotations tools. IBM's MPEG-7 annotation tool is a commercial, MS Windows based system to annotate atomic video sequences encoded in MPEG-1 or MPEG-2 [8]. The annotations are based on static descriptors and stored locally. Similarly, an annotation of MPEG audio and video files with MPEG-7 metadata can be obtained with IBM's Multimodal Annotation Tool [9]. The commercial Ricoh Movie Tool [15] allows the creation of spatio-temporal MPEG-7 annotations. It stores descriptions of the video structure of individually selectable tags locally. Caliph&Emir is an open source MPEG-7 based photo annotation and retrieval tool written in Java [13]. It offers substantial options for photo related metadata annotations. Storage and retrieval takes place on the file system. FilmEd is a Windows based system that allows the collaborative indexing, browsing, annotation, and discussion of video content between multiple groups at remote locations [17]. The resulting annotations are shared via a database server in a proprietary metadata format which is a combination of MPEG-7 and DC. The web-based COALA project contains a LogCreator that allows the creation of MPEG-7 video descriptions [6]. It has a predefined structure specially designed for the annotation of TV news. VIDETO of ZGDV is an MPEG-7 tool that has been specially designed for TV news annotation [20]. It contains automatic scene boundary detection for a temporal decomposition of MPEG-1 videos by freely customizable

description sets. The resulting metadata description is stored locally. Probably the most comprehensive automatic parser and generator of DC metadata for online resources is DescribeThis [3]. The web-based tool allows an automatic creation of DC from HTTP, FTP, or ISBN content. Similar approaches can be found at the Nordic DC metadata creator [4], or DC-dot web-sites [2]. The so far mentioned tools mostly stress on compliance to metadata standards, but don't facilitate - in general - the community wide exchange of multimedia artifacts. This deficit is tackled in the web-based system American Authors. It has been specially designed to support literature studies at the MIT and aims at being a digital resource for engaging media artifacts with literary texts [11, 16]. Additionally, it's a protected community environment that requires authentication. The annotated files are stored compliant to DC and Text Encoding Initiative [19]. Similar arguments also hold for BSCW [12] and Wikipedia [21], with the exception of compliance to the previously mentioned metadata standards.

Summarizing, none of the existing systems covers web collaboration on multimedia artifacts of various data types with a platform independent system by simultaneously providing compliance to MPEG-7 and DC. Additionally, there is only one project dealing with the needs of cultural science communities in particular. However, that project doesn't offer compliance to MPEG-7, which is needed for more elaborated descriptions of time based media. We will now introduce MEDINA and explain its features to foster collaboration and KM in cultural science communities.

#### 4 MEDINA

The MPEG-7 Encoding of Dublin Core Information and Naming Application (MEDINA) has been specially designed for cultural science communities of researchers and students within the collaborative research center "Media and Cultural Communication". In particular, it targets multidisciplinary research community in a project of the film studies. The community is physically distributed in the department of film studies at the Ruhr-Universität Bochum, the center in Cologne, University Bonn, and Munich. Its members have diverse backgrounds of education, e.g. film studies, history of art, graphical design and are on diverse levels of profession, i.e. full professors, research assistants, and students.

Fig. 1 shows the user interface of MEDINA. It is divided into two parts. On the left hand side, the tabbed pane shows the user groups in MEDINA. Depending on the individual access rights the user can navigate through and add media to the corresponding media sets. On the right hand side, the tabbed pane offers four different functionalities: Thumb-preview of the media sets including a media-player, refinement of DC annotation by more sophisticated MPEG-7 descriptors, the MPEG-7 preview of the resulting file, and a temporal decomposition window containing the decompositions of a medium. MEDINA has been

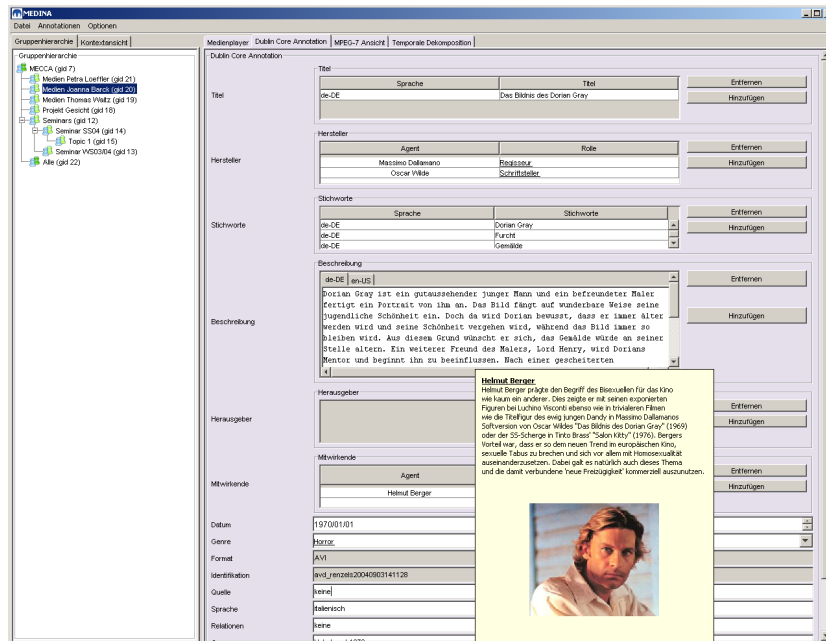
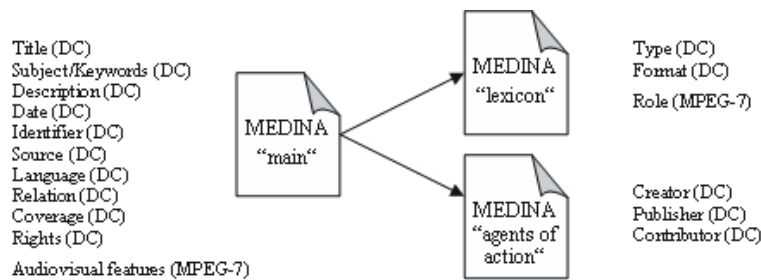


Figure 1: MEDINA user interface when uploading an DC document (in German)

implemented as a platform independent Java application and offers two options to annotate multimedia artifacts. First, a new multimedia annotation can be created from scratch. Second, a DC annotation might be uploaded for an automated conversion to MPEG-7. Fig. 1 shows an uploaded DC document for a subsequent manual refinement possible with MPEG-7 descriptors. Technically, the conversion from DC to MPEG-7 is a two-stage process. First, the DC document is parsed internally in order to create a “normalized” DC document that contains the 15 elements (or a subset, in case not all of them have been used in the DC document) of the DC Metadata Element Set. The second step is the conversion itself. Here, the mapping between the DC Metadata Element Set [1] and more sophisticated descriptors of the MPEG-7 multimedia metadata standard takes place. It is similar as proposed in [7], but has been refined for the particular needs of cultural science communities (cf. Appendix). For the sake of usability, additional MPEG-7 metadata concerning technical aspects of the medium such as visual encoding, frame rate, color space, etc. are extracted automatically.

The DC elements are used for a loose metadata description offering a lot of semantic freedom in the scope of a more general metadata element. Thus, we’ve subdivided the metadata file into three documents. A MEDINA “main” MPEG-7 metadata file with descriptors for technical information about the medium as well as most of the DC information. In addition, it contains references to two other MPEG-7 descriptions. The first type is a reference to “agents

of action”. The second type is a reference to terms defined in MPEG-7 classification schemes that are used as a multi language “lexicon”. Thus, we are able to combine the more sophisticated role descriptors of MPEG-7 with DC elements. Fig. 2 shows the previously described structural decomposition of MPEG-7 documents in MEDINA. On the server side, we’ve set up an eXist-DB [14] for the storage of the MPEG-7 metadata. Additionally, there is an affiliated FTP-server. The FTP-server is used for automated up- and download of multimedia artifacts by the community to the common repository.



**Figure 2:** Structural decomposition of MPEG-7 files in MEDINA

## 5 Conclusions & Outlook

In this paper we presented MEDINA for collaboration and KM in multimedia repositories of cultural science communities. The use of MEDINA demonstrated the usefulness of multimedia description standards for transparent and flexible collaboration in open and scalable repositories. Even more, the maturity of collaboration via KM systems is increased by the given compatibility to the MPEG-7 and the DC standard. Due to its platform independence and an easy to use user interface, cultural scientists can now collaborate anytime and anywhere by exchanging sets of multimedia artifacts, their temporal decompositions, and their annotations about them via a common repository.

Current research has two goals. On the one hand side, we will embed (semi-) automatic feature extraction tools into MEDINA. In this aspect, we plan to retrieve spatial information from the key frame analysis of videos for a further spatio-temporal video analysis. In addition, we integrate algorithms for color and shape segmentation of digital images as well as term analysis in text documents. On the other hand, we have recently set up an MPEG-7 community with our colleagues from the University of Klagenfurt and the Know Center in Graz. Thus, we intend to create testbeds of MPEG-7 metadata in order to facilitate conformance tests of MPEG-7 applications. In this aspect, MEDINA may serve as (semi-) automatic generator of MPEG-7 metadata.

## Acknowledgements

This work was supported by German National Science Foundation (DFG) within the collaborative research centers SFB/FK 427 “Media and Cultural Communication”, the SFB 476 “IMPROVE”, and by the 6<sup>th</sup> Framework IST programme of the EC through the Network of Excellence in Professional Learning (PROLEARN) IST-2003-507310. We’d like to thank our colleagues for the inspiring discussions. In particular, we’d like to thank our student workers Dominik Renzel and Monika Pienkos.

## References

1. ANSI/NISO Z39.85-2001: The Dublin Core Metadata Element Set. <http://www.niso.org/standards/resources/Z39-85.pdf>, September 10 2001.
2. DCdot - Dublin Core metadata editor. <http://www.ukoln.ac.uk/metadata/dcdot/>, August 3 2000.
3. DCS - Sand’s Dublin Core Service. <http://www.describethis.com/>, 2004.
4. Dublin Core Metadata Template. <http://www.lub.lu.se/cgi-bin/nmdc.pl>, 2003.
5. Dublin Core Metadata Initiative. <http://dublincore.org/>, 2005.
6. N. Fatemi, O. A. Khaled, and G. Cory. La recherche dinformation dans les archives du téléjournal à la tsr. *Flash informatique*, (4):8 – 11, 2003.
7. J. Hunter. An Application Profile which combines Dublin Core and MPEG-7 Metadata Terms for Simple Video Description. [http://www.metadata.net/harmony/video\\_appln\\_profile.html](http://www.metadata.net/harmony/video_appln_profile.html), Feb. 12 2002.
8. IBM MPEG-7 Annotation Tool. <http://www.alphaworks.ibm.com/tech/videoannex>, April 10 2003.
9. IBM Multimodal Annotation Tool. <http://www.alphaworks.ibm.com/tech/multimodalannotation>, August 9 2002.
10. ISO. Information technology – Multimedia content description interface – Part 8: Extraction and use of MPEG-7 descriptions. Technical Report ISO/IEC TR 15938-8:2002(E), ISO, 2002.
11. W. Kelley. American authors. <http://metaphor.mit.edu/projects/americanauthors.html>, 2004.
12. K. Klöckner. BSCW - Educational Servers and Services on the WWW. In *Proc. of the International C4-ICDE Conf. on Distance Education and Open Learning “Competition, Collaboration, Continuity, Change”*, Adelaide, September 9-14, 2000.
13. M. Lux, W. Klieber, and W. Granitzer. Caliph & Emir. In *Proceedings of the 19<sup>th</sup> International C4-ICDE Conf. on Distance Education and Open Learning “Competition, Collaboration, Continuity, Change”*, Berlin, Germany, November 7-10, 2004. <http://www.codata.org/04conf/abstracts/DataVis/Lux-SemanticsinMultimediaRetrievalandAnnotation.htm>.
14. W. Meier. eXist: An Open Source Native XML Database. In A. B. Chaudhri, M. Jeckle, E. Rahm, and R. Unland, editors, *Web, Web-Services, and Database Systems, NODe 2002 Web and Database-Related Workshops, Erfurt, Germany, October 7-10, 2002, Revised Papers*, volume 2593 of LNCS, Springer-Verlag, Berlin Heidelberg, pages 169 – 183, 2003.
15. Ricoh Movie Tool Home. <http://www.ricoh.co.jp/src/multimedia/MovieTool/>, September 10 2004.
16. L. Ridgway. MetaMedia Translates into Classroom Innovation. *IS&T: News about Information Services and Technology throughout MIT*, 19(5), May/June 2004.
17. R. Schroeter, J. Hunter, and D. Kosovic. FilmEd - Collaborative Video Indexing, Annotation and Discussion Tools Over Broadband Networks. In *Proceedings of the 10<sup>th</sup> Intl. Multimedia Modelling Conference, Brisbane, Australia, January 5-7*, pages 346–353, 2004.

18. C. P. Snow. *The Two Cultures*. Cambridge University Press, Cambridge, 1959.
19. TEI - Text Encoding Initiative. <http://www.tei-c.org/>, September 6 2003.
20. VIDETO - Video Description Tool.  
[http://www.zgdv.de/zgdv/departments/zr1/Produkte/videto/index.html\\_en](http://www.zgdv.de/zgdv/departments/zr1/Produkte/videto/index.html_en),  
May 5 2004.
21. J. Wagstaff. Wikipedia: It's Wicked. *The Wall Street Journal Online, Loose Wire*.  
<http://www.theproduct.com/6m105/readings/spring04/encyclopedia.pdf>, 2004.

## Appendix: Mapping between Dublin Core and MPEG 7

dc:contributor	Reference:	/Mpeg7/DescriptionUnit/MultimediaContent //CreationInformation/Creation /Creator[Role/Name="Contributor"] /AgentRef
	Referenced Element:	/Mpeg7/Semantics/Agent/Name
		In MEDINA we use different documents for multimedia and agents. Contributors are considered as agents with the role "Contributor". Agent descriptions are referenced from multimedia descriptions. The role is defined as a term in an MPEG-7 classification scheme.
dc:coverage		/Mpeg7/DescriptionUnit/MultimediaContent//CreationInformation/CreationCoordinates/{Location,Date}
		In MEDINA we consider the combination of Location and Date as the coverage of a multimedia description.
dc:creator	Reference:	/Mpeg7/DescriptionUnit/MultimediaContent //CreationInformation/Creation /Creator[Role/Name="Creator"]/AgentRef
	Referenced Element:	/Mpeg7/Semantics/Agent/Name
		In MEDINA we use different documents for multimedia and agents. Creators are considered as agents with the role "Creator". Agent descriptions are referenced from multimedia descriptions. The role is defined as a term in an MPEG-7 classification scheme.
dc:date		/Mpeg7/DescriptionUnit/MultimediaContent//CreationInformation/Classification/Date
		The MPEG-7 standard defines the same encoding for a date, so Date is compliant with both MPEG-7 and Dublin Core.
dc:description	Abstract or FreeText	/Mpeg7/DescriptionUnit /MultimediaContent//CreationInformation /Creation/Abstract/FreeTextAnnotation
	Account:	Reference to graphical URI of a JPG image, that is located on representation: a central FTP server or a local file system.
		In MEDINA, we define three of the above mentioned descriptions: abstract, free-text account and reference to graphical representation.
dc:format	Reference to Term representing Format:	/Mpeg7/DescriptionUnit/MultimediaContent //MediaInformation/MediaProfile/MediaFormat /Content/@href
	Referenced Term:	/Mpeg7/DescriptionUnit /ClassificationScheme/Term
		In MEDINA the format is defined as a reference to a predefined term of the corresponding MPEG-7 Classification Scheme.
dc:identifier		/Mpeg7//MultimediaContent//MediaInformation /MediaIdentification/EntityIdentifier/Mpeg7/DescriptionUnit//*/@id
		In MEDINA we need two identifiers for multimedia content. The first one is used like in Dublin Core to capture the media's identifier from a formal identification system such as DOI, URI, or ISBN in case it exists. The second one is the identifier of a medium within our system.

dc:language	/Mpeg7/DescriptionUnit/MultimediaContent //CreationInformation/Classification/Language
dc:publisher	Reference: /Mpeg7/DescriptionUnit/MultimediaContent //CreationInformation/Creation /Creator[Role/Name="Publisher"]/AgentRef Referenced Element: /Mpeg7/Semantics/Agent/Name  In MEDINA we use different documents for multimedia and agents. Publishers are considered as agents with the role "Publisher". Agent descriptions are referenced from multimedia descriptions. The role is defined as a term in an MPEG-7 classification scheme.
dc:relation	Node for this multimedia description: /Mpeg7/DescriptionUnit/Relationships/Node[./@id="THISID" ...] Source Node: /Mpeg7/DescriptionUnit/Relationships/Node[./@id="ID" ...] Corresponding Relation: /Mpeg7/DescriptionUnit/Relationships/Relation[./@source="THIS" and ./@target="ID" and ./@type="<SOME_RELATION>"]  An MPEG-7 Node is defined as a reference to any other MPEG-7 object. An MPEG-7 Relationships description contains a node for this multimedia description and for all related nodes along with corresponding relations, where the node for this Description nodes along with a related node is target. The type of the relation is a reference to a term, which is defined in an MPEG-7 Classification Scheme.
dc:rights	/Mpeg7/DescriptionUnit/MultimediaContent//CreationInformation/Creation/CopyrightString  In MEDINA, this descriptor is represented by a string, since MEDINA uses the rights management via the LAS Server Security Manager.
dc:source	Node for this multimedia description: /Mpeg7/DescriptionUnit/Relationships/Node[./@id="THISID" ...] Source Node: /Mpeg7/DescriptionUnit/Relationships/Node[./@id="ID" ...] Corresponding Relation: /Mpeg7/DescriptionUnit/Relationships/Relation[./@source="ID" and ./@target="THISID" and ./@type="dc:source"]  An MPEG-7 Node is defined as a reference to any other MPEG-7 object. An MPEG-7 Relationships description contains a node for this description and for all related nodes along with multimedia corresponding relations, where the node for this Description is source and a related node is target. The type of the relation is a reference to a term, which is defined in an MPEG-7 Classification Scheme.
dc:subject	/Mpeg7/DescriptionUnit/MultimediaContent//CreationInformation/Classification/Subject/Keyword
dc:title	/Mpeg7/DescriptionUnit/MultimediaContent//CreationInformation/Creation/Title  In MEDINA we use the xml:lang attribute of Title for i18n.
dc:type	Term Reference: /Mpeg7/DescriptionUnit/MultimediaContent//CreationInformation/Classification/Genre/@href Referenced Term: /Mpeg7/DescriptionUnit/ClassificationScheme/Term  In MEDINA we use MPEG-7 classification schemes for the definition of terms. These terms can defined in multiple languages. We use the xml:lang attribute of Term for i18n.