



MEDIA ENGINEERING WITH A CONTENT HUB SYSTEM. GENERIC CONTENT TRANSFORMATION FOR MOBILE MEDIA

Michael A. Herzog ^{1,2)}, Matthias Trier ²⁾

¹⁾ University of Applied Sciences FHTW Berlin, 10313 Berlin, herzog@fhtw-berlin.de,
<http://inka.fhtw-berlin.de/herzog>

²⁾ TU Berlin, trier@sysedv.tu-berlin.de, <http://sysedv.tu-berlin.de>

Abstract: *A primary problem in multimedia content production is the plethora of formats and the resulting multiple point-to-point transformation processes in different content formatting specifications. However, these content transformation processes can be conceptually broken down into a limited set of abstract principles. This allows a unified approach that theoretically enables the transformation of any given input channel into any desired output. Based on this principle, the present article introduces the concept of Generic Content Transformation and the related project MOCCA, which develops the corresponding software implementation of generic content transformation for a variety of input and output channels. To balance technical similarities with specific properties of different media channels, we suggest a combination of an XML-based structure (MocML) and a high level authoring environment. Generic Content Transformation further allows for novel means of accessing assets via compound search algorithms. In the second part we summarize and discuss current results and experiences in three scenarios of media production, including mobile media.*

Keywords: *Multimedia Authoring, Mobile Content, Cross Media Production, Content Engineering, Media Retrieval*

1. MOTIVATION

The following article describes a technical approach driven by a process management perspective on the field of media authoring automation. Its application in mobile content aggregation was first introduced 2007 at IDAACS conference. This contribution is a technically extended version of the proceedings paper [1].

To support different multimedia channels in multimedia applications it is usually necessary to transcode content for various presenting environments in a manual and time consuming process. Regularly, media authors have to work with specialized authoring systems or content management solutions. Occasionally they have the opportunity to automatically distribute content from a single original source to different mobile channels with a variety of screen resolutions, aspect ratios, operating systems and media data formats. But that requires exactly predefined situations. The access to and reuse of multimedia content depends on well-organized data structures with high quality metadata. The corresponding manual tagging is usually only done in a few larger companies with enhanced

XML-workflows, where industrialized processes of media production are established, like in the Printing Industry [2].

Activities like storing, marking, and archiving multimedia content in databases and repositories is currently required by applications in Enterprise Content Management, Information Lifecycle Management, and Data Mining. The most articulated objective is the concentration of enterprise data with the help of unified media formats [2]. In the area of annotating metadata to text-based information, the evolution of software development has been spurred by the introduction of XML and other technology standards. Further impact resulted from recent significant advances of semantic systems. But also in the domain of image data and time-based media a consolidation of formats is beneficial. Here, the addition of metadata, automated detection of meta-information from multimedia data (e.g. [3]), machine learning methods [4], support with semantic technologies [5], a significant increase in media availability, and shorter search operations improved results. Coinciding with these technical developments, enterprises need to improve their efficiency in coping with the increasing volume of

media content, which has usually grown between 60 and 200 percent annually [6]. To manage this situation, they adopt technologies that enable content centralization, structuring, and automated combination. Despite this explicit requirement many multimedia solutions still prosper, although they are designed for very individual tasks and problems, or create new data formats and codecs. The existing multitude of project data and distribution formats seems unlimited. This digital "impedance mismatch" of media content makes it hard or impossible to exchange the results of creative design work and to manage changes across the heterogeneous format requirements of different authoring tools. The availability of interfaces is mostly limited to exchanging proprietary package formats and a sparse selection of dedicated assets. Asset import is often restricted to low complexity and oriented towards uniform codecs. Media production processes require significant time and manpower for conversion and extraction of content into compatible file formats, both activities that do not add value to the work. A loss of creative components, like layout, navigation and functionality is very likely.

These general issues in the context of multimedia authoring can also be applied to the special scenario that is analyzed and discussed in the mobile learning case study of this paper: teachers often need to produce learning media material for mobile use without any support if they want to offer it to their students. They tend to prefer their current authoring environments like Powerpoint to produce such media types. They also want to combine different media content from former lectures. At present, there is no efficient way to generate multiple mobile multimedia content formats based on that kind of application. On the other hand a lot of time is required to find the right media assets and other data, to compare versions and variants, in order to achieve high quality results from hard disc or external archives. Thus a lecturer today has to master all challenges of a multimedia author.

2. RESEARCH OBJECTIVE

The previous section indicated that a primary problem in the domain of content production is caused by the plethora of formats and the resulting multiple point-to-point transformation processes in different content formatting specifications. However, when these transformation processes are analyzed, it can be recognized that transformation of authored multimedia content packages in practice can conceptually be broken down to a small set of abstract principles. This assumption enables a unified approach that theoretically allows the transformation of any given input channel into any

desired output. Based on these principles, a project called MOCCA was formed to develop a software-based approach to generic content transformation for a variety of input and output channels. Technical similarities and specific properties of different channels, their influence on an abstract process of content transformation and an XML-based structure to store transformational data represent some core work packages.

A related objective was to take advantage from a unified data description of authored media content in repositories. Data management with a centralized enterprise repository allows accessing unified media content with advanced semantic technologies. The semantic junction of text information and media data can provide a quick and easy way to retrieve content assets without explicit metadata. In combination with content based retrieval (CBR) methods and automatic media crawling this enables a comprehensive media data mining solution for enterprise application provided via GNU Public License (GPL).

Faced with the research objective of developing a sustainable software platform for transforming a variety of input formats into several output channels, the project aims to overcome current point-to-point integration facilities and to evolve towards a framework, that allows for the subsequent addition of further transformations between other input and output formats. This challenge necessitated the definition of a generic set of architectural structures and processes to deal with the import, editing, and export of multiple formats. A further issue is storing all formats in a unified project format. These considerations motivated and initiated the work on a method for Generic Content Transformation. It was developed to allow transcoding data from entire media containers to alternative forms of representation. The vision is to develop a transformation framework offering the modularity and flexibility of a Content HUB both as a desktop application as well as a background Enterprise Content Integration application.

This contribution begins by outlining the research intention of developing a sustainable framework that allows the addition of new input and output channels to a generic content hub. Sample implementations for various input and output channels in different domains are presented to illustrate the approach. In the subsequent section, the underlying concept of Generic Content Transformation is summarized and finally the envisioned and resulting benefits of employing the MOCCA framework for Generic Content Transformation are discussed.

3. THE CONCEPT OF GENERIC CONTENT TRANSFORMATION (GCT)

With the increasing replacement of proprietary data representation by transparent XML descriptions, the conversion of content can be improved to a large extent. For this objective, a concept of Generic Content Transformation (GCT) was developed that enables transcoding of entire media container formats to alternative representation models. GCT utilizes a generalization of these media containers with the intention of independent and standard-dependent data storage and transformation (figure 1). Further, a corresponding product model for transformation, storage and aggregation of media content from different authoring sources was defined.

The GCT concept is based on the assumption that most authoring systems have modeled their project data using the same principles, including the following aspects:

- (1) Text and media assets are stored frame based in connection with vector representation;
- (2) A limited number of functions for navigation and interaction is used;
- (3) Time-based media assets are embedded with basic control functions;
- (4) Other functionality is mostly represented in textual code structures, stored as software routines like scripts.

The transformation of these aspects to an independent XML representation can be automated with import plugin components. The arbitrary transfer of aspects (1) to (3) to other representations for distribution purposes is only restricted by limitations of output containers. The resulting output presentation will get an identical layout and full navigation functionality.

The user operation for transcoding a project data package with text, pictures and navigation elements requires a simple three step process: import the project content, select a transformer, and specify the project folder name within the file system or database repository. In a next step, the user has access to a variety of authoring functions or to an immediate output without changes to a variety of destination formats.

For complex or inaccessible programming that relates to aspect (4) the project's application provides visualizations of extracted code at an optional authoring panel so that a content developer is encouraged to add the functionality based on this documentation. In cases of exceptions (text overflows or unexpected instructions), comprehensive editing and authoring functions are available.

The GCT concept was informed by a preliminary

technical study of 22 selected project data formats from authoring systems. It was found that more than 80% of all project information can be stored and reprocessed within a complexity-reduced and unified data structure without special effort. The remaining 20% of project information, e.g. special application logic or control of proprietary media containers are detectable and transformable, but usually need a significant amount of time for analysis and implementation.

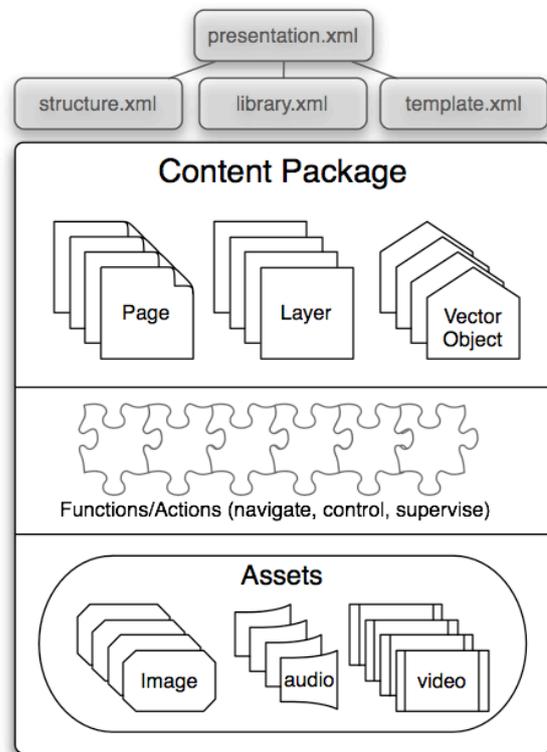


Fig. 1: General content package description

The use of corresponding individual solutions in the highly dynamic environment of changing format specifications requires an adequate amount of resources to get a 95%+ transformation supported for a long time period. The GCT concept was not designed to solve these very specific tasks of logic conversion. It is, however, capable to integrate and embed other software. The GCT approach is an efficiency driven model based on the 80/20 paradigm. Not all functions and codecs have to be transformed into all distribution formats. The objective is to keep nearly all components in a generic storage format with an acceptable cost-benefit-ratio.

To handle the complexity of a large software architecture, the GCT concept utilizes a Content HUB product model based on separated modules interoperating on specific data channels (figure 3). It includes an authoring application that works on MocML as an authoring container format, a wide variety of plugins for importing and distributing

other project formats, an optional database repository mirroring the full MocML-structure of each project dataset, a separate crawler engine collecting compatible project data formats from enterprise file directories and transforming them automatically into MocML, and a standalone retrieval application to give semantic access to the multimedia data repository.

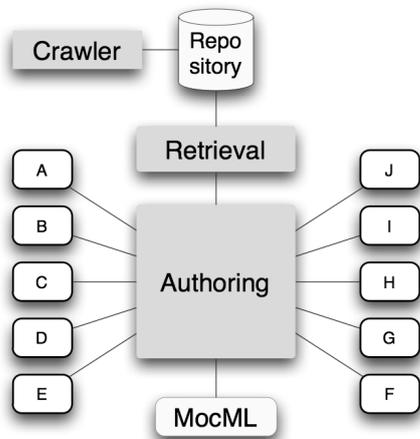


Fig. 2. Content HUB architectural concept

4. APPLICATIONS, EXPERIENCE AND EVALUATION

The concept of GCT with its principles of transformation, the unified data description, and the MocML format have been tested and evaluated in several applications. Three scenarios are discussed now, namely (1) converting MHP to Flash in a television-related media environment, (2) converting Powerpoint to SCORM compliant learning objects in a conventional e-learning scenario, and (3) providing several mobile media formats like Podcasts or MP3 in a mobile learning scenario. The applications were part of an iterative approach to validating and extending the software implementation approach of the GCT concept.

In the first scenario of the GCT project, a transformation and authoring solution for a branch of the primary German public television company ARD-Digital¹ was developed with the intention to bring MHP-Content² to Adobe Flash containers while using a separate XML-based data format, where all data issues could be represented. The transformation process is fully automated, but at the authoring layer, changes of layout, text and pictures are possible. The authoring environment follows the principles of multimedia authoring architectures described in [7] and is includes some context driven features discussed in [8].

Another requirement in this first scenario was a further distribution channel in order to experience initial advantages from the generic approach. Accordingly, the system provides various templates to arrange MMS messages from all imported content and to export SMIL data directly, compliant to a lot of mobile format issues [9].

The independent MocML data representation was realized with RelaxNG³ and is closely aligned with the SVG 1.2 W3C Standard⁴. It is geared and in some points restricted to the requirements of import and export demands, especially to interactive TV documents and their structural issues [10]. The selection of a lean media description format was largely motivated by efficiency criteria. Alternative formats like OOXML with more than 6,500 pages of specification have not been considered for the implementation and evaluation of the GCT media transformation, authoring, and retrieval approach. However, more complex standardized media descriptions should stay objects of investigation. A general description of the resulting runtime data structure is shown in figure 3.

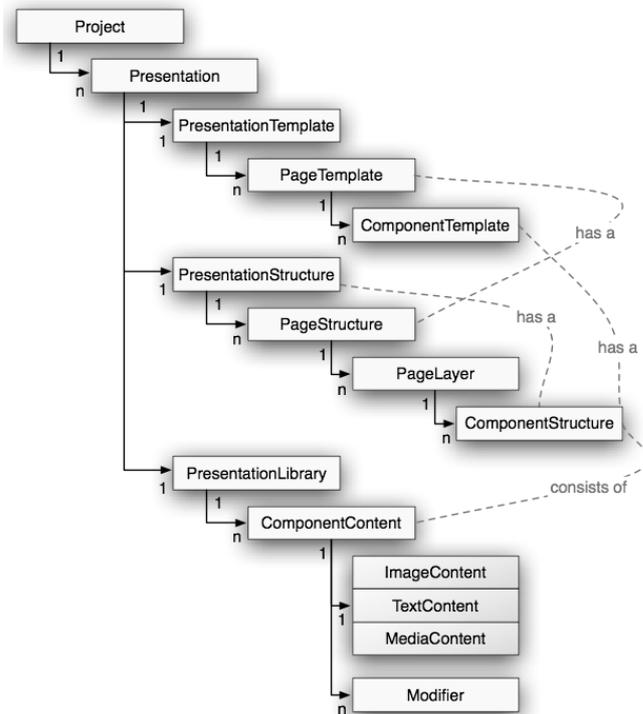


Fig. 3. Mocca OO runtime data tree (extract)

The transformation, authoring, and retrieval solution is implemented with Java in combination with the Open Source community frameworks GEF⁵, JMF⁶, OpenLaszlo⁷. It uses the open development

¹ ARD Digital: <http://www.ard-digital.de>
² MHP Multimedia Home Platform: <http://www.mhp.org/>

³ REgular LAnguage for XML Next Generation, <http://relaxng.org/>
⁴ Scalable Vector Graphics, <http://www.w3.org/Graphics/SVG/>
⁵ Graphical Editing Library, <http://gef.tigris.org/>
⁶ Java Media Framework, <http://java.sun.com/products/java-media/jmf/>
⁷ <http://www.openlaszlo.org/>

platform Eclipse RCP⁸.

A second scenario to evaluate the feasibility and potential of GCT concerned E-Learning production processes involving SCORM creation and mobile learning applications. The objective is to show that the GCT concept improves the process efficiency by 30 to 50 percent, in some cases more. The extension of the MocML data description format to handle the complexity of authored containers is another objective of the project. Such channel-independent data representation provides much potential, e.g. for improving information retrieval across the boundaries of different content package formats.

The implementation started with the extension of input and output connectors for E-Learning authoring purposes and audio editing facilities. The work is practically related to the practical and organizational production process of interactive video- and audio based instructional design content at the vbc.studiolab⁹ for the Virtual Global University¹⁰. There, a dedicated XML based production process was established which uses PowerPoint as an authoring tool to produce all content except audio and video data [11]. The PowerPoint material is to be transformed, combined with related audio and video, converted to HTML and aggregated with templates to arrive at complete SCORM packages.

A lean and more efficient version of this fully automated SCORM production process is available as an ODP2CP online service¹¹.

The extension of the Java-based MOCCA application with more import and output facilities is a current working package as well as the realization of the repository with the automated Data Crawler and the Retrieval Component with advanced CBR features [12]. A summary of the extended architecture resulting from implementing scenario 2 is shown in figure 4. The most central component of architecture is the *authoring application*. It has to meet the needs of fully automated GCT as well as of traditional multimedia authoring. The requirement of frequent manual content customization was satisfied by elaborate media processing and authoring editors. To import and reduce complexity of authoring container formats several *import plugin* components for transformation to MocML and integration of media assets like MPEG4- and QuickTime were built. To realize *export plugins* a number of open XML-Frameworks for Media Processing were used.

i.e. OpenLaszlo¹² for Flash applications, QuickTime for Java¹³ for Podcasts and universal Quicktime export, and Cocoon¹⁴ for PDF, XHTML and SCORM processing. The framework is not particularly focused on creating mobile content but offers a template-concept to be open in content adoption for most conceivable purposes including mobile distribution. With predefined template sets it is flexible enough and easy to use for producing mobile media content.

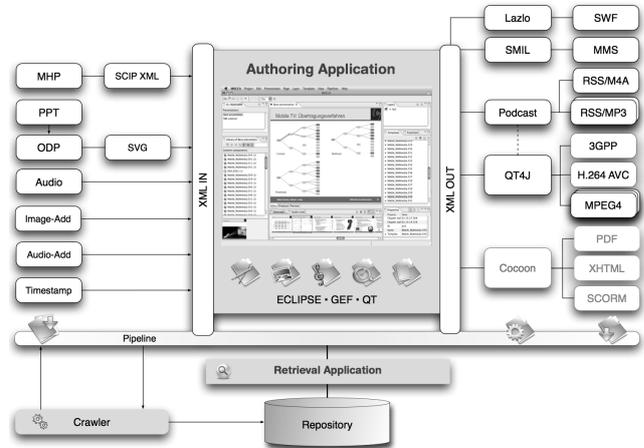


Fig. 4: Software Architecture MOCCA 1.2 (12/2007)



Fig. 5: M-Learning Podcast on iPod

As an example of how these results could also encourage the support of innovative and novel content output formats, mobile learning media production has been examined as the third scenario. The objective is to establish a simple production process to complement a physical academic lecture with a mobile podcast distribution (cf. figure 5).

A related survey about »Podcasting as a distribution channel for M-Learning« conducted

⁸ <http://www.eclipse.org/>

⁹ Video based content studio laboratory at University of Technology Berlin: <http://www.sysedv.tu-berlin.de/E-Learning>

¹⁰ Virtual Global University: <http://www.v-gu.de>

¹¹ Web application for online transformation from Open Document Presentation (ODP) to SCORM documents, available at <http://www.moccaonline.de>

¹² <http://www.openlaszlo.org/>

¹³ <http://developer.apple.com/quicktime/qtjava/>

¹⁴ <http://cocoon.apache.org/>

between October 2005 and February 2006 at the University of Applied Sciences FHTW Berlin confirmed that considerable value of a presentation of lectures with audio and slide pictures via podcasts (for a more extensive account of the results, cf. [13]). The study showed a preference of podcast media over written material (like textbooks) for learning purposes with 70% of respondents interested in continuously using podcast media and a mobile usage rate of 68%.

However, the production process requires too much manual work, especially when teachers produce it themselves. The survey identified that a wide variety of students use different MP3 players and applications at their computers. For a maximum of acceptance it was necessary to distribute at least two versions of podcasts (figure 6): A segmented MP3 file with the linked slides collected with an RSS XML file and a version of M4A, where slides and audio elements are connected in an MPEG4 container. A very simple mp3 audio distribution with only one chapter per lecture was also used in the first semesters but later seemed to have become obsolete for most students.

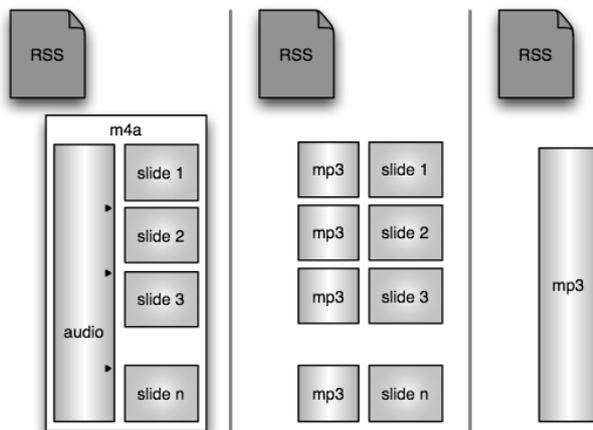


Fig. 6: Distributed RSS Podcast Variants

An elaborate part of GCT development was the implementation of audio authoring facilities, which were mainly based on the Quicktime for Java Framework. With the integration of audio and slides in the MOCCA framework and the addition of a multichannel podcast export, it became possible to decrease production time dramatically from 4 times to 0.5 times of a sequentially recorded lecture. This means a reduction of 800% of process duration, which was currently evaluated in the postgraduate course »Mobile Computing«¹⁵.

To get an advanced access to multi-channel media material, a repository database was implemented based on the GCT-concept. In that

process, the complete document structure of MocML was matched to a relational scheme (fig. 7). The MOCCA *Data Crawler* collects compatible data from specified file hierarchies and converts it to the MocML format using the content HUB import plugins. The *Database Importer* migrates MocML files to the repository database and creates CBR feature indexes for image or audio assets. Because of performance reasons CBR indices are separately held in text files.

Based on this repository an advanced and flexible search over all collected data elements becomes possible. The retrieval client application can be used from every place in a company to search for media content. Using the text information from related authored media material, single (non-textual) assets become visible without having been tagged manually before.

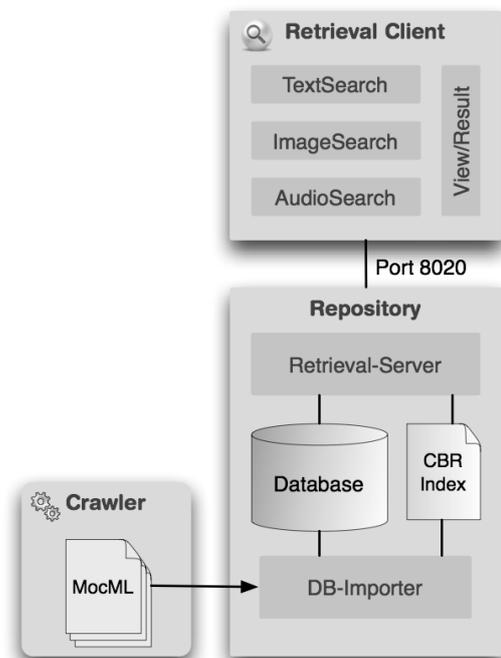


Fig. 7: MOCCA Retrieval System

In addition, similar assets are detected by CBR features. The implemented methods are *Color Layout* and *Edge Histogram* for images or *Mel Frequency Cepstrum Coefficients* (MFCC) for audio content. Among other benefits, these features are particularly effective in finding higher resolutions of an image resource, finding similar graphic design patterns, or finding lessons of the same lecturer. The employed search technology is based on Apache Lucene and was extended with novel proprietary CBR modules. Most enhancements resulted from blending text and content based retrieval techniques with a unified information access to the repository.

The project's current focus is the development of functionality for aggregating authored multimedia

¹⁵ Postgraduate course Mobile Computing at University of Applied Sciences FHTW Berlin. <http://inka.fhtw-berlin.de/moco>

content from strongly differing media production processes. The expected result is genuine cross media authoring for diversified media container formats with special emphasis on output for mobile media.

5. CONCLUSION

Our experiences from surveys and the current software development show: The approach of a generic Content HUB for different areas in content generation and aggregation is feasible and beneficial. The proposed strategy of integrating different connected media types enables synergistic effects, including the following benefits:

Archiving and information retrieval. With transparent access to different content container variants and versions a huge base of information with special links to media elements can be exploited for specific search operations. For example: The search request "Find all pictures from pages with the phrase 'iceberg'" needs only a quick and simple algorithm to detect the phrase and find media assets connected with this page. The ability to search and browse across all collected multimedia presentations, converted to the independent XML format, and the semantic links among assets and pages or layout patterns gives much more efficient search results than a global search through documents, especially when there is no metadata. Often criticized CBR methods are efficient enough to raise retrieval quality perceptibly when they get combined with other methods in a sequential data mining approach.

Persistence and compliance. Concepts for long term archiving and accessing of container-based data may integrate the MocML meta container format. It will thus not be necessary to support the semantics of all possible proprietary formats. The described content transformation is proposed to be an integrated part of the global enterprise backup strategy and meets the needs of compliance demands in the area of media container formats.

Shared production. Merging of authoring results from different applications and the centralized integration and management of content parts opens a way to new work processes. An ad hoc parallelization of different media authoring processes based on different applications will be possible.

Parallel media distribution. The content HUB we introduced provides the benefit of classic content management systems: producing one content in different distribution formats. Moreover, the direct output of content parts into a CMS compatible XML format is a manageable part of development.

Change management. The transformation of

proprietary container formats into XML based data representations and their management using repositories increases transparency. With the help of a cross referencing ID method it will become possible to identify and manage changes in correlated containers.

User empowerment. By providing a simple transformation process, a low complexity interface, and a selection of templates, our research indicates that even inexperienced users are empowered to distribute multiple variants of content with a single desktop application. In the case of podcast creation for mobile learning this means that a teacher imports the slides and the recorded audio from his lesson, combines it with a prerecorded trailer and exports it with three clicks to different podcast versions. A flash animation export using a navigation template could be an easily attainable output.

Process optimization. The GCT concept introduced above was developed to obtain a substantial enhancement of process efficiency ratios in the area of multimedia production. Process simplification, reductions in process cycle time and editing time are particularly addressed with MOCCA's Pipeline Concept and its Template Constructor. As an integrated functional module, this combination described above enables dedicated channels of transformation where it is possible to transform an import container into a distribution format without any authoring interaction automatically in the background. Thus, by using the concept of General Content Transformation and the related software implementation MOCCA, every additional integration step of a dedicated pipeline instance becomes a step forward to achieve the Content HUB vision.

ACKNOWLEDGMENT

Research and development was partially supported by the project IKAROS of German Federation of Industrial Cooperative Research Associations (AiF) and by grants from Federal Ministry of Education and Research (BMBF). We would like to thank our colleagues in Berlin, Dortmund, and Görlitz for their useful discussions and suggestions. Special thanks are due to Daniel Möller, Andreas Hohendorf, Björn Hoffmann, Roman Frohn, Johannes Stein and the PPA-Media-Repository-Group for their extensive development activities in the MOCCA project.

REFERENCES

- [1] M.A. Herzog, M. Trier "Multimedia Production with a Content HUB Environment. Media Engineering for Mobile Computing Courseware" *IEEE Proceedings of conference*

- "Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS'2007)"*, Dortmund, Germany. Sept. 2007
- [2] A. Benlian, C. Grau, T. Hess, Y. Braunstein, "Dissemination of Content Reutilization Practices in the German and US Book Publishing Industry", in: *Journal of Media Business Studies (JOMBS)*, 3. Jg. 2006, No. 2, p. 41 - 61
- [3] M. Engelhardt, A. Hildebrand, T.C. Schmidt, "A Semantic Approach to Automated Content Augmentation for eLearning Objects", Michael Auer, Ursula Auer (Eds.) *Proceedings of the ICL 2005*, Kassel University Press, Villach (Austria), 2005
- [4] K. Franke, K.R. Müller, B. Nickolay, R. Schäfer: Pattern Recognition, *Proceedings of 28th DAGM Symposium*, Berlin, Germany, September 12-14, 2006, Springer 2006
- [5] A. Hotho, A. Nürnberger, G. Paaß, "A Brief Survey of Text Mining." *German journal for computer linguistics and speech technology*, 20(1):19-62, ISSN 0175-1336, 2005.
- [6] Survey 4/2005, <http://www.jupiterresearch.com>
- [7] D.C. Bulterman, L. Hardman, "Structured multimedia authoring", *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)*, v.1 n.1, p.89-109, February 2005, revised Version of an 1993 article
- [8] A. Scherp, S. Boll, "Context-driven smart authoring of multimedia content with xSMART", *Proceedings of the 13th annual ACM international conference on Multimedia*. ACM Press 2005.
- [9] D.C. Bultermann, L. Rutledge, "SMIL 2.0 - Interactive Multimedia for Web and Mobile Devices", Springer, 2004
- [10] R. Goularte, E. Moreira, M. Pimentel, "Structuring interactive TV documents", *Proceedings of the 2003 ACM symposium on Document engineering*, ACM Press New York 2003, p. 42-51
- [11] C. Müller, M. Trier, M.A. Herzog, "Process-oriented Production of Learning Units for sustainable E-Learning Offerings", M. Breitner, G. Hoppe (editors), *E-Learning. Einsatzkonzepte und Geschäftsmodelle*, *Proceedings of ELWH*, Physica/Springer Berlin 2005.
- [12] M.S. Lew, N. Sebe, C. Djeraba, R. Jain, "Content-Based Multimedia Information Retrieval: State of the Art and Challenges" *ACM Transactions on Multimedia Computing, Communications and Applications*, Vol. 2, No. 1, February 2006, p. 1-19.
- [13] J. Sieck, A. Fiedler, M.A. Herzog: Mobile Information Systems and Mobile Learning. In: Is information technology shaping the future of higher education? *Proceedings of 12th International Conference of European University Information Systems*, p. 347-353, Tartu, Estonia, 2006



Michael A. Herzog, teaches and researches in Media Technology, Content Engineering and IT supported Business Process Management at University of Applied Sciences FHTW Berlin and TU Berlin. He has worked for about 15 years as senior executive and founder of several IT-enterprises in Berlin (Germany) and Treviso (Northern Italy) concerning media technology and software development.

As conference chair he is responsible for the series »Wireless Communication and Information« (WCI), »Culture and Computer Science«, and »Media Production Berlin« (MP). In 2004 he received the "Multimedia-Award of the University of Potsdam" for a component-based E-Learning Content Production approach developed for the Virtual Global University. Since 2006 he is Research Fellow of Alcatel-Lucent-Foundation for Communication Research.



Matthias Trier, teaches and researches at TU Berlin. He is Head of IKM Research Group of the Institute for Business Informatics. The Group develops innovations for Knowledge Management and E-Learning and has published in international top journals. His PhD work focused on new methods for the visualization

and analysis of electronic knowledge communities. Further, he is a practitioner working on efficient Video- and XML-based production processes of E-Learning contents and received the "Multimedia-Award of the University of Potsdam in 2004. He is further responsible for the Berlin Branch of the internationally active Virtual Global University VGU.